





CHICAGO CIRCLE



LIBRARY

This book is the gift of

Dreyfus Fund

UNIVERSITY of ILLINOIS





Q

1

A773

Y.1

1948

PER



AUSTRALIAN JOURNAL  
OF  
SCIENTIFIC RESEARCH

SERIES A  
PHYSICAL SCIENCES

VOLUME 1

MELBOURNE  
1948



# AUSTRALIAN JOURNAL OF SCIENTIFIC RESEARCH

Published by the Council for Scientific and Industrial Research  
in collaboration with the Australian National Research Council

Issued in two Series :

Series A – Physical Sciences

(Four issues each year appearing March, June, September, December)

Series B – Biological Sciences

(Four issues each year appearing February, May, August, November)

Price : Each series 30/- per annum, separate issues 7/6 each

---

## EDITORIAL BOARD

Chairman and Editor : N. S. Noble, D.Sc.Agr.

Members :     Professor W. J. Dakin, D.Sc.  
                  Professor E. J. Hartung, D.Sc.  
                  Professor L. H. Martin, Ph.D.  
                  Professor J. G. Wood, Ph.D., D.Sc.

---

All enquiries and manuscripts should be forwarded to :

The Editor,  
Australian Journal of Scientific Research,  
C/o Council for Scientific and Industrial Research,  
314 Albert Street, East Melbourne, C.2, Victoria.

Printed by Australasian Medical Publishing Co. Ltd., Sydney.



## CONTENTS

NUMBER 1, MARCH 1948

	PAGE
Diffraction of Light from Sources of Finite Dimensions. By G. H. Godfrey .. .. .	1
The Angular Distribution of D-D Neutrons scattered by Deuterons. By J. F. Darby and J. B. Swan .. .. .	18
The Influence of a Horizontal Wall on the Motion of a Falling Oil Drop. By V. D. Hopper and Alison M. Grant .. .. .	28
Atmospheric Pressure Changes: The Importance of Deviations from the Balanced (Gradient) Wind. By C. H. B. Priestley .. .. .	41
Observations on the Variable Source of Cosmic Radio Frequency Radiation in the Constellation of Cygnus. By J. G. Bolton and G. J. Stanley ..	58
The Deformation and Recrystallization of an Alloy containing Two Phases. By R. W. K. Honeycombe and W. Boas .. .. .	70
Lubrication of Metal Surfaces by Silicone Films. By J. N. Gregory and Marjorie J. Newing .. .. .	85
"One- " and "Two-dimensional" Partition Chromatographic Separations of Organic Acids on an Inert Sheet Support. By J. W. H. Lugg and B. T. Overell .. .. .	98
Studies on the Lignin of <i>Eucalyptus regnans</i> . I. Thiolignin. By F. N. Lahey and J. W. T. Merewether .. .. .	112



## NUMBER 2, JUNE 1948

	PAGE
The Theory of Symmetrical Crossed Flexure Pivots. By W. H. Wittrick	121
Relaxation Methods Applied to Two Problems of Two-Dimensional Stress Distribution Involving Mixed Boundary Conditions. By W. H. Wittrick and W. Howard .. .. .	135
Local Isotropy in the Turbulent Wake of a Cylinder. By A. A. Townsend	161
The Velocity of Sound in Gases. By R. L. Abbey and G. E. Barlow	175
The Disordering of $\beta$ Brass by Cold Work. By R. W. K. Honeycombe and W. Boas .. .. .	190
The Rate of Evaporation of Water through Duplex Films. By A. R. Gilby and E. Heymann .. .. .	197
Equilibria in Furfural-Water Systems under Increased Pressure and the Influence of Added Salts upon the Mutual Solubilities of Furfural and Water. By R. G. Curtis and H. H. Hatt .. .. .	213
Alkylene Trithiocarbonates and Attempted Syntheses of 2, 3-Dimercaptopropanol (B.A.L.). By C. C. J. Culvenor and W. Davies.. ..	236
Studies on the Lignin of <i>Eucalyptus regnans</i> . II. The Nature of the Hydroxyl Groups and the Presence of the Carbonyl Group in Thiolignin. By J. W. T. Merewether .. .. .	241



## NUMBER 3, SEPTEMBER 1948

	PAGE
The Scattering of Fast Positrons and Electrons in Gold (for the Range 0.39–1.1 Mev.). By W. B. Lasich .. .. .	249
Cosmic Ray Measurements made during the Cruise of H.M.A.S. <i>Wyatt Earp</i> 1947–48. By D. E. Caro, P. G. Law, and H. D. Rathgeber ..	261
Hydrogen Atmospheres in the Absence of Thermodynamic Equilibrium. I. The Populations of the Upper Atomic States. By R. G. Giovanelli	275
Hydrogen Atmospheres in the Absence of Thermodynamic Equilibrium. II. The Populations of the Lower Atomic States. By R. G. Giovanelli	289
Hydrogen Atmospheres in the Absence of Thermodynamic Equilibrium. III. The Emission of Radiation. By R. G. Giovanelli .. .. .	305
Statistical Mechanics of High Polymer Solutions. By A. R. Miller ..	319
N-Acyl Derivatives of Aromatic Ketimines. By J. E. Banfield, G. M. Brown, F. H. Davey, W. Davies, and T. H. Ramsay .. .. .	330
The Thermal Decomposition of Synthetic and Natural Alunite: An Investigation by X-ray Diffraction, Electron Diffraction, and Electron Microscope Methods. By N. S. Bayliss, J. M. Cowley, J. L. Farrant and G. L. Miles .. .. .	343

## NUMBER 4, DECEMBER 1948

	PAGE
Plane Waves in an Ionized Gas with Static Electric and Magnetic Fields Present. By V. A. Bailey .. .. .	351
Hydrogen Atmospheres in the Absence of Thermodynamic Equilibrium. IV. The Solar Chromosphere. By R. G. Giovanelli .. ..	360
The Electronic Charge and the Oil Drop Method. By V. D. Hopper ..	369
Some Aspects of the Glow Discharge Between Coaxial Cylinders in the Presence of a Non-Homogeneous Axial Magnetic Field. By J. M. Somerville, K. S. W. Champion, and E. K. Bigg .. ..	400
Translation and Development in Two Dimensional Fields with Special Reference to Pressure Variations. By R. W. James .. ..	412
The Influence of Vertical Ionic Drift on a "Chapman Region". By C. B. Kirkpatrick .. .. .	423
Radio Superrefraction in the Coastal Regions of Australia. By F. J. Kerr	443
A New Interferometer. By E. R. Johnson and J. F. M. Scholes ..	464
The Visible Absorption Spectrum of Bromine in Solution. By N. S. Bayliss, A. R. H. Cole, and B. G. Green .. ..	472
The Free Energies of Hydration of Gaseous Ions. By N. S. Hush ..	480
Index to Volume 1 .. .. .	495









## *Foreword*

*The Council for Scientific and Industrial Research, with the approval and co-operation of the Australian National Research Council, ventures to launch this new periodical, the Australian Journal of Scientific Research.*

*The issue will at first be in two series covering papers relating to physical sciences and biological sciences, respectively.*

*The aim is to encourage scientific endeavour in the Commonwealth and it is believed that both quality and quantity of material available for publication will justify the step now taken.*





# DIFFRACTION OF LIGHT FROM SOURCES OF FINITE DIMENSIONS

By G. H. GODFREY\*

[Accepted for Publication January 19, 1948]

## *Summary*

The distribution of light intensity in the diffraction pattern of continuous light sources is treated analytically for Fraunhofer diffraction phenomena. Sources and diffracting apertures of any dimensions are first discussed, and a result of general application is derived which relates the illumination and dimensions of the diffraction pattern to the dimensions of the source and aperture. A source consisting of a luminous strip is considered in more detail, the form of the pattern is derived for such a source, and tables are supplied to facilitate numerical evaluation. Consideration is given to the maxima and minima of intensity in the image, and the conditions under which these are observable are indicated. In particular, attention is drawn to certain diffraction effects which do not seem to have been previously recorded and which are in agreement with the theory. These effects consist of faint but visually discernible fringes associated with the image of a strip when viewed through a narrow slit. Applications to the optics of the telescope and spectroscope are discussed.

## I. INTRODUCTION

The diffraction of light from a point source has been exhaustively treated in the literature of the subject for many types of diffracting aperture, and the extension to the case of a linear source of infinitesimal width has also been made for the more important types of aperture. Struve(1) has dealt with a semi-infinite source bounded by a straight line when the diffraction aperture is circular. His treatment is restricted to the Fraunhofer type of diffraction phenomena, and he gives a table from which the numerical value of the illumination at any point in the image plane may be obtained. Several investigators have dealt with a source in the form of a straight strip, but only to the extent that it enters into some particular problem such as the resolving power of a spectrometer. The following analysis gives some results of a general nature, applying to sources and apertures of any shape. This is followed by a detailed treatment of the strip source for rectangular and circular apertures.

The discussion is restricted to Fraunhofer diffraction phenomena in which the diffraction pattern is examined in the plane of the image, and to the case where the source is of uniform brightness. It is assumed that the points comprising the source are non-coherent. The treatment is based on Huyghens's principle, as interpreted by Kirchhoff.

The results contained in the following section are of general interest but are referred to only incidentally in the work which follows it.

\* Division of Physics, C.S.I.R. (part-time).

## II. SOURCE OF ANY SHAPE

## 1. EFFECT ON THE DIFFRACTION PATTERN OF CHANGES IN THE DIMENSION OF APERTURE AND SOURCE

In any optical system, using a small monochromatic source of area  $dS$  at a fixed distance from the aperture, the intensity of the diffraction pattern at an arbitrary origin of coordinates in the image plane may be written in the form

$$dI = K.f(x,y,a,b)dS \quad \dots\dots\dots (1)$$

where  $a$  and  $b$  are parameters of the nature of scale factors, depending on the dimensions of the aperture,  $x$  and  $y$  are the coordinates of the geometric image of the element  $dS$ , and  $K$  is a constant. We will assume that  $a$  and  $b$  refer to dimensions parallel to the  $x$  and  $y$  axes respectively. From the well-known theory of optical systems the dimensions of the geometrical image are proportional to those of the source for a fixed position of the latter. Thus, if the linear magnification is  $k$  and we write  $C = k^2 K$ , the illumination at the origin for an extended source is

$$I = C \iint f(x,y,a,b) dx dy \quad \dots\dots\dots (2)$$

the integration extending over the area of the geometric image.

For a point source there is the following general property of the diffraction pattern, due to Bridge and quoted by Rayleigh(2). *If the dimensions of the aperture parallel to two fixed directions at right angles be multiplied by  $1/m$  and  $1/n$  respectively the dimensions of the pattern in the same directions are multiplied by  $m$  and  $n$  and the illumination at corresponding points is multiplied by  $1/m^2 n^2$ .*

Referring to equation (1) this property is equivalent to the statement

$$f(x,y,a,b) = m^2 n^2 f(\bar{x}, \bar{y}, a/m, b/n)$$

where  $\bar{x} = mx$  and  $\bar{y} = ny$ .

Substituting this result in (2),

$$I = mnC \iint f(\bar{x}, \bar{y}, a/m, b/n) d\bar{x} d\bar{y}$$

where the integration extends over the image of a source corresponding to the original source but with linear dimensions parallel to the  $x$  and  $y$  axes multiplied by the factors  $m$  and  $n$  respectively. This result may be written  $I' = I/mn$ , the interpretation of which leads to the following general theorem.

*If the dimensions of the source parallel to any two directions at right-angles be increased by the factors  $m$  and  $n$  respectively, and the corresponding dimensions of the aperture altered inversely, then the corresponding dimensions of the diffraction pattern are increased by the factors  $m$  and  $n$  but the brightness at corresponding points is reduced in the ratio  $1/mn$ .* This result, of course, satisfies the energy principle.

Examples of this principle in particular cases will be obvious in what follows. It will be seen that in general if a small source can be used with a given aperture to approximate to a point source in producing diffraction effects, an equally good approximation results, but with reduced illumination, if a large source is used with a proportionately reduced aperture.

One particular application of the general principle may be mentioned at this stage. The dispersion of a diffraction grating depends on the relative displacement for different wavelengths of corresponding principal maxima from



a common central maximum. Thus, if all the dimensions of the grating (including width of rulings) are altered by the factor  $m$ , and those of the slit in the inverse ratio, the displacements referred to, and hence also the dispersion, will alter by the factor  $1/m$ . Moreover, all points in the diffraction patterns for any two wavelengths which previously coincided will continue to coincide and the light distributions in the patterns remain geometrically similar, so that the resolving power is unaltered. The intensity of the spectrum, however, will be altered in the ratio of  $m:1$ . It should be noted that the above result agrees with the ordinary rule for a slit of infinitesimal width, viz. that the resolving power depends on the total number of rulings, and the dispersion on their spacing.

## 2. COMPLEMENTARY SOURCES

When the illumination at any point of the diffraction pattern is known for any given source of uniform brightness, then it is also known for a second source which together with the original source would constitute an infinitely extended source. It is necessary merely to subtract the illumination in the former case from that due to an infinitely extended source of the same brightness. For example, when the diffraction pattern due to a source consisting of a luminous thick wire of uniform brightness is known, it is also known for an extended source close to which the same wire in a non-luminous condition is placed so as to intercept part of the light.

## III. UNIFORM STRIP SOURCE

The diffraction pattern for a source in the form of a long straight luminous strip of uniform width will now be considered for points in the plane of the image close to the axis of the optic system.

### 1. GENERAL CONSIDERATIONS

#### (a) *Formula for the Illumination*

For a strip source it is more convenient to employ the results for a linear source than to utilize equation (1), which applies to a point source.

In Figure 1,  $P$  is a point of a luminous strip, perpendicular to the plane of the paper and of width  $\tau$ , and  $Q$  is the geometric image of  $P$  produced by an optical system  $AB$ . Choose  $x$  and  $y$  coordinates in the image plane, the  $y$  axis being perpendicular to the length of the strip and  $x$  being the distance of  $Q$  from an arbitrary origin of coordinates  $O$ ; and let the geometric images of the edges of the strip be represented by  $x=X$  and  $x=X+t$ . The contribution to the illumination at  $O$  due to an elementary strip at  $P$  of width  $d\tau$  is given by

$$dI = Kf(x)d\tau \dots\dots\dots (3)$$

where  $K$  depends on the brightness and length of the source and  $f(x)$  on the shape of the aperture. Since  $dx = kd\tau$ , where  $k$  is the magnification, the illumination at  $O$  due to the complete strip may be written

$$(I_x)_t = C \int_{X'}^{X+t} f(x)dx \dots\dots\dots (4)$$

where  $C$  is equal to  $K/k$ . The distance  $X$  is always measured from the edge of the image indicated in the figure and is to be considered positive when  $O$  is to the left of that edge.

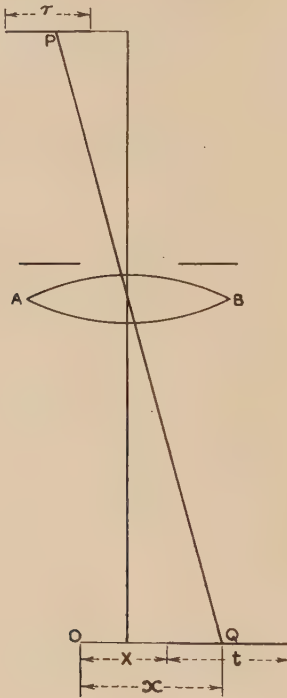


Fig. 1

(b) *Illumination at the Centre of the Image for a Symmetrical Aperture*

For points in the plane of the image close to the axis of the optic system, the diffraction pattern for an aperture symmetrical with regard to the strip source is symmetrical about the centre line of the geometric image, and hence the illumination along this line is a maximum or minimum. The conditions for a maximum or minimum in this case are now analysed.

At the centre of the image  $x = -t/2$ , and we have from (4)

$$\begin{aligned} d(I_X)_t/dX &= C[f(X+t) - f(X)] \\ [d(I_X)_t/dX]_{X=-t/2} &= C f(t/2) - f(-t/2) \\ &= 0 \end{aligned}$$

since  $f(x)$  is essentially an even function from the symmetrical nature of the pattern. Also,

$$d^2(I_X)_t/dX^2 = C[d\{f(X+t)\}/dX - d\{f(X)\}/dX],$$

so that  $[d^2(I_X)_t/dX^2]_{X=-t/2} = 2C[d\{f(x)\}/dx]_{x=t/2}$

since  $df(x)/dx$  is an odd function.

Now the diffraction pattern for the infinitely fine line source as given by (3) has a series of maxima and minima, i.e.  $df(x)/dx$  is zero for certain values of  $x$ .



Between any two consecutive values of  $x$  for which this occurs the quantity  $df(x)/dx$  is either negative everywhere or positive everywhere. For the values of  $t$  that make the value of  $df(x)/dx$  at  $x=t/2$  negative, the illumination at the centre of the geometric image of the strip source is, of course, a maximum, and for those values that make it positive the illumination is a minimum. This result may be stated in general terms as follows:

*For any aperture which is symmetrical with regard to the centre line of a strip source, the centre of the diffraction pattern has a local maximum of intensity for semi-widths of the geometric image between 0 and  $t_1$ ,  $t_2$  and  $t_3$ , etc., and a local minimum for semi-widths between  $t_1$  and  $t_2$ ,  $t_3$  and  $t_4$ , etc.;  $t_1$ ,  $t_3$ , etc. represent the distances of the successive minima from the centre of the corresponding diffraction pattern for a straight line source, and  $t_2$ ,  $t_4$ , etc. the distances of the successive maxima.*

### (c) Application to Particular Apertures

Circular and rectangular apertures are of the most frequent occurrence in optical measurements and therefore these will be discussed in detail. Certain general considerations, however, apply to both.

If the length of the rectangle is parallel to the strip,  $f(x)$  takes the same form in both cases, viz.

$$f(x) = A\varphi(px) \quad \dots\dots\dots (5)$$

where

$$p = \pi a' / \lambda' D' = \pi \alpha' / \lambda' = \pi a / k \lambda D \quad \dots\dots\dots (6)$$

$A$  is a constant depending on the dimensions of the aperture, and  $a$  is the width of the rectangular aperture or the diameter of the circular aperture as the case may be, both being located in the object space.  $D$  is the distance of the aperture from the strip,  $\lambda$  the wavelength in the object space, and  $k$  the linear magnification produced by the optic system. The quantities  $a'$  and  $D'$  are the images of  $a$  and  $D$  respectively,  $\alpha'$  is the angle subtended by  $a'$  at the image, and  $\lambda'$  the wavelength in the image space.

Thus for both these cases (4) may be written in the form

$$(I_X)_t = C' \int_{pX}^{p(X+t)} \varphi(x) dx,$$

where  $C' = A.C/p$ ,

$$\text{or} \quad (I_X)_t = C' \int_{\bar{X}}^{\bar{X} + \bar{t}} \varphi(x) dx \quad \dots\dots\dots (7)$$

$$\text{where} \quad X = \bar{X}/p \text{ and } t = \bar{t}/p \quad \dots\dots\dots (8)$$

We shall now derive an expression for the constant  $C'$ . Consider the illumination on the image plane due to a source extending to infinity in all directions. As this source may be regarded as equivalent to two continuous sources extending respectively from 0 to  $\infty$  and from 0 to  $-\infty$  we have, for the resultant intensity  $I_m$ , in virtue of the principle of complementary patterns mentioned above,

$$(I_X)_\infty + (I_{-X})_{-\infty} = I_m \quad \dots\dots\dots (9)$$

so that

$$(I_0)_\infty = I_m/2 \quad \dots\dots\dots (10)$$

which gives the value of the illumination at the edge of the geometric image of a semi-infinite sheet.

From (7)

$$C' = I_m / \left[ 2 \int_0^\infty \varphi(x) dx \right] \quad .$$

Now equation (7) shows that if  $\bar{X}$  and  $\bar{t}$  be kept constant the corresponding value of the illumination  $(I_X)_t$  is proportional to  $C'$  or  $I_m$ , i.e. to the area of the aperture. But to maintain  $\bar{X}$  and  $\bar{t}$  constant,  $X$ ,  $t$ , and hence also the actual strip width  $\tau$ , must by (8) vary inversely as  $p$ , i.e. as the linear dimensions of the aperture. Hence if the dimensions of the aperture be multiplied by a factor  $m$ , and those of the strip divided by  $m$ , the dimensions of the diffraction pattern are divided by  $m$ . This accords with the general principle enunciated in Section II.

For the sake of convenience, however,  $I_m$  will be taken as unity, so that

$$C' = 1 / \left[ 2 \int_0^\infty \varphi(x) dx \right] \quad \dots\dots\dots (11)$$

Writing

$$I_{\bar{X}} = C' \int_0^{\bar{x}} \varphi(x) dx \quad \dots\dots\dots (12)$$

(7) gives

$$(I_X)_t = I_{\bar{X}+\bar{t}} - I_{\bar{X}} \quad \dots\dots\dots (13)$$

and since  $\varphi(x)$  is essentially an even function

$$I_{-\bar{X}} = -I_{\bar{X}} \quad \dots\dots\dots (14)$$

If the function of  $I_{\bar{X}}$  be tabulated for a sufficiently large number of values of  $\bar{X}$ , then the value of the illumination at any point outside or inside the geometric image of the strip may be obtained from the tables by means of (13) and (14).

Alternatively, if tables of  $(I_X)_\infty$  are available the illumination may be obtained from the identity

$$(I_X)_t = (I_X)_\infty - (I_{X+t})_\infty \quad \dots\dots\dots (15)$$

together with (9). Rectangular and circular apertures will now be considered separately.

## 2. DIFFRACTION BY A RECTANGULAR APERTURE

### (a) General Formula

For a rectangular aperture(3)

$$\varphi(x) = [(\sin x)/x]^2 \quad \dots\dots\dots (16)$$

so that

$$\int_0^{\bar{X}} \varphi(x) dx = \text{Si}(2\bar{X}) - (\sin^2 \bar{X})/\bar{X}$$

where

$$\text{Si}(2X) = \int_0^{2X} [(\sin x)/x] dx.$$

Thus

$$\int_0^\infty \varphi(x) dx = \pi/2,$$

and (11) gives

$$C' = 1/\pi,$$

whence by (12)

$$I_{\bar{X}} = [\text{Si}(2\bar{X}) - (\sin^2 \bar{X})/\bar{X}]/\pi \quad \dots\dots\dots (17)$$

Tables of the sine integral  $\int_0^X [(\sin x)/x]dx$  are available(4) so that the

function  $I_{\bar{X}}$  may be calculated from (17). Values of  $I_{\bar{X}}$  for values of  $\bar{X}$  from zero to 34.5 are given in Table 1. Values of  $(I_X)_t$  may be found directly from the figures tabulated by the use of equations (13) and (14).

TABLE 1  
VALUES OF  $10^5 \cdot I_{\bar{X}} = 10^5 [\text{si}(2\bar{X}) - \sin^2 \bar{X}]/\bar{X}]/\pi$

$\bar{X}$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	00000	03180	06338	09455	12509	15482	18356	21115	23744	26230
1	28563	30735	32741	34575	36238	37730	39055	40217	41223	42081
2	42806	43403	43886	44267	44559	44775	44928	45030	45092	45124
3	45138	45141	45141	45145	45157	45182	45222	45279	45353	45444
4	45550	45671	45803	45944	46091	46241	46390	46537	46678	46810
5	46933	47044	47142	47228	47299	47358	47405	47440	47464	47481
6	47490	47495	47497	47497	47498	47500	47505	47514	47528	47548
7	47573	47604	47640	47681	47726	47774	47825	47877	47930	47982
8	48031	48079	48122	48162	48197	48220	48254	48275	48291	48303
9	48311	48316	48319	48320	48321	48321	48321	48323	48326	48332
10	48340	48351	48364	48381	48400	48421	48444	48469	48495	48521
11	48548	48574	48599	48622	48644	48664	48681	48696	48708	48718
12	48725	48730	48734	48736	48737	48737	48750	48763	48775	48787
13	48799	48810	48821	48832	48842	48852	48862	48872	48881	48890
14	48899	48908	48916	48924	48932	48940	48947	48954	48961	48968
	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5
15	48975	49006	49034	49058	49080	49100	49119	49138	49154	49171
20	49189	49205	49223	49241	49260	49279	49297	49317	49334	49352
25	49367	49382	49397	49411	49421	49431	49439	49449	49455	49465
30	49470	49479	49485	49494	49500	49505	49515	49520	49525	49535

(b) *Maximum and Minimum Intensities at the Centre*

The illumination at the centre of the geometric image is, from (13) and (14),

$$\begin{aligned} (I_{-t/2})_t &= I_{\bar{t}/2} - I_{-\bar{t}/2} \\ &= 2I_{\bar{t}/2} \quad \dots\dots\dots (18) \end{aligned}$$

and is immediately obtainable from the tables for different values of  $t$ . Now we have seen that for this central illumination to be a minimum or a maximum the expression  $[df(x)/dx]_{x=t/2}$  should be respectively positive or negative. Here  $f(x)$  is given by (5) and (16) and has minimum (zero) values for  $p = \pi, 2\pi, 3\pi$ , etc. The values of  $p$  corresponding to successive maximum values of  $f(x)$  are well



known from Schwerd's calculations(5), the first few being  $p=0, 1.430\pi, 2.459\pi, 3.471\pi$ . Thus for widths of geometrical image between the ranges of  $\bar{t}$  from 0 to  $2\pi$ ,  $2.860\pi$  to  $4\pi$ ,  $4.918\pi$  to  $6\pi$ , etc., the illumination at the centre is greater than that in its immediate surroundings (i.e. it is a local maximum), and for values of  $\bar{t}$  between  $2\pi$  and  $2.860\pi$ ,  $4\pi$  and  $4.918\pi$ , etc., the central illumination is a local minimum. Thus we should expect a somewhat pronounced minimum of intensity for a width in the vicinity of  $\bar{t}=2.430\pi$ . It is therefore of interest to trace, by means of Table 1, the variations in intensity for a width of image corresponding to, say,  $\bar{t}=2.40\pi$ , as we pass outward from the centre. Table 2 gives these variations for a change in  $\bar{X}$  from  $-1.2\pi$  to  $+1.3\pi$  and also the corresponding variations for the complementary case.

TABLE 2

VALUES OF  $(I_X)_t = I_{\bar{X}+\bar{t}} - I_{\bar{X}}$  FOR  $\bar{t}=2.4\pi$  AND CORRESPONDING VALUES OF  $1-(I_X)_t$ , RECTANGULAR APERTURE

$\bar{X}$	$(I_X)_t$	$1-(I_X)_t$	$\bar{X}$	$(I_X)_t$	$1-(I_X)_t$
$-1.2\pi$	0.90660	0.09340	$+0.1\pi$	0.38073	0.61927
$-1.1\pi$	0.90810	0.09190	$+0.2\pi$	0.28971	0.71029
$-1.0\pi$	0.91233	0.08767	$+0.3\pi$	0.20991	0.79009
$-0.9\pi$	0.91660	0.08340	$+0.4\pi$	0.14498	0.85502
$-0.8\pi$	0.91763	0.08237	$+0.5\pi$	0.09646	0.90354
$-0.7\pi$	0.91140	0.08860	$+0.6\pi$	0.06367	0.93633
$-0.6\pi$	0.89382	0.10618	$+0.7\pi$	0.04443	0.95557
$-0.5\pi$	0.86163	0.13847	$+0.8\pi$	0.03544	0.96456
$-0.4\pi$	0.81293	0.18707	$+0.9\pi$	0.03283	0.96717
$-0.3\pi$	0.74737	0.25263	$+1.0\pi$	0.03318	0.96682
$-0.2\pi$	0.66687	0.33313	$+1.1\pi$	0.03375	0.96625
$-0.1\pi$	0.57533	0.42467	$+1.2\pi$	0.03293	0.96707
0.0	0.47793	0.52207	$+1.3\pi$	0.03035	0.96965

It will be noticed that the illumination at the central minimum (which occurs at  $\bar{X} = -\bar{t}/2 = -1.2\pi$ ) is less by about 1.2 per cent. than that at a point whose distance from the centre of the pattern is one-sixth of the image width ( $\bar{X} = 0.8\pi$ ), a contrast which may well be visually perceptible.

To derive the conditions for observing this minimum, suppose  $\tau$  to be the actual width of the strip source. The angle subtended by the strip at the aperture is, from equation (6), given by

$$\theta = \tau/D = \bar{t}/(pkD) = \lambda \bar{t}/\pi a \quad \dots\dots\dots (19)$$

The angular width  $\theta$  of the strip for the minimum at  $\bar{t}=2.4\pi$  is obtainable from (19) for any aperture of width  $a$ . For an aperture of 0.5 mm. and  $\lambda$  equal to  $6 \times 10^{-5}$  cm., the value is about 10 minutes.

The visual observation of this minimum does not seem to have been previously recorded. It may be observed as follows:

Place a strip of white paper about an inch wide against a black background and view it at a distance of about 8.3 metres through a slit of width 0.5 mm. held close to the eye with its edges parallel to the strip. A faint but definite dark fringe will be discernible down the centre of the strip.

The complementary case of a source, continuous except for a central strip of the same width as that of the strip source just considered, is also of interest. Table 2 shows that the central maximum existing here presents a contrast with its immediate surroundings about ten times as great as that between the central minimum and its surroundings in the case just considered. However, the absolute value of the illumination over this minimum is very low compared with that of the region beyond the edge of the geometric image of the strip, and the illumination needs to be reasonably great for easy observation. A window bar viewed against a bright sky or even a distant chimney stack is quite suitable for observing the bright central line, the slit being adjusted according to the angle subtended by the object.

(c) *Non-Central Maxima and Minima*

We shall now consider non-central maxima and minima. In this connection Table 3, for a strip width given by  $\bar{t}=3.4\pi$ , is of interest. It shows a minimum

TABLE 3

VALUES OF  $(I_X)_t = I_{\bar{X}+\bar{t}} - I_{\bar{X}}$  FOR  $\bar{t}=3.4\pi$  AND CORRESPONDING VALUES OF  $1-(I_X)_t$  FOR RECTANGULAR APERTURE

$\bar{X}$	$(I_X)_t$	$1-(I_X)_t$	$\bar{X}$	$(I_X)_t$	$1-(I_X)_t$
$-1.7\pi$	0.94516	0.05484	$-0.8\pi$	0.92909	0.07091
$-1.6\pi$	0.94387	0.05613	$-0.7\pi$	0.92105	0.07895
$-1.5\pi$	0.94042	0.05958	$-0.6\pi$	0.90247	0.09753
$-1.4\pi$	0.93568	0.06412	$-0.5\pi$	0.86990	0.13010
$-1.3\pi$	0.93154	0.06846	$-0.4\pi$	0.82816	0.17184
$-1.2\pi$	0.92881	0.07119	$-0.3\pi$	0.75557	0.24443
$-1.1\pi$	0.92820	0.07180	$-0.2\pi$	0.67482	0.32518
$-1.0\pi$	0.92938	0.07062	$-0.1\pi$	0.57277	0.42723
$-0.9\pi$	0.93067	0.06933	0.0	0.48463	0.51537

for the bright strip at the point approximately given by  $\bar{X} = -1.1\pi$ , about one-third of the width of the strip from its centre, with, of course, a corresponding maximum in the complementary case of a black strip against a bright background. The minimum for the corresponding bright strip is too faint to be observable. The intensity of the maximum in the case of the black strip is about 3.5 per cent. greater than the intensity at the point  $\bar{X} = -0.9\pi$  (about

half-way between the centre and edge of the strip) so that the contrast is quite sufficient for observation. By symmetry there is a corresponding maximum at an equal distance on the other side of the centre. If a suitable dark object against a bright background be viewed through a slit, and the width of the latter adjusted to give the central bright line, then, on opening the slit about half as wide again (corresponding to increase in  $\bar{t}$  from  $2.4\pi$  to  $3.4\pi$ ), the two bright maxima in the position indicated can be observed replacing the central bright line.

Returning to Table 2, it will be observed that the illumination due to a bright strip of the width indicated falls off continuously in the region outside the geometric image (bounded by  $\bar{X}=0$ ) till a point  $\bar{X}=0.9\pi$  is reached. It

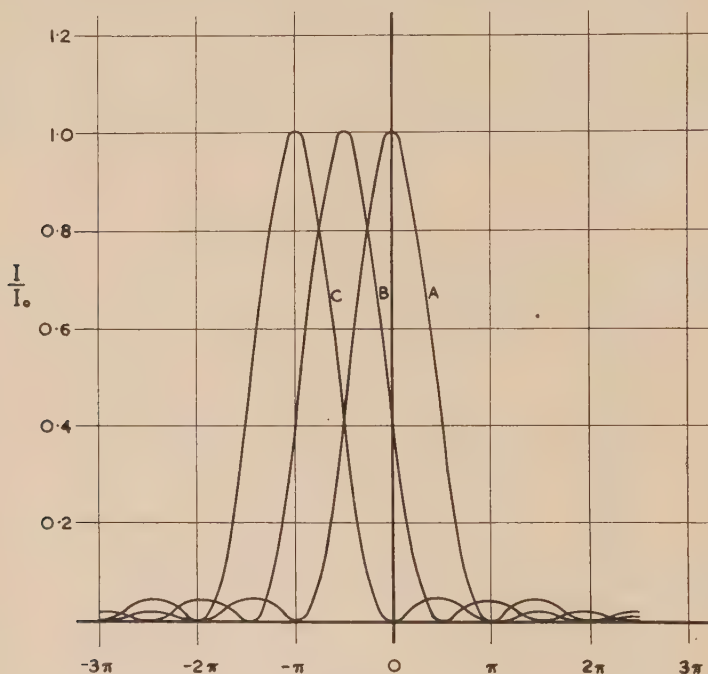


Fig. 2

then increases to a sharp maximum near  $\bar{X}=1.1\pi$ , at which point its value is about 3 per cent. greater than at the point  $\bar{X}=0.9\pi$ . The fact that these two points are approximately equally spaced from the point  $\bar{X}=\pi$  is significant, and raises the question of the position of maxima and minima in general. On differentiating (13) we have from (12)

$$d(I_X)_t/dX = C'[\varphi(\bar{X}+\bar{t}) - \varphi(\bar{X})].$$

Thus from (16) on substituting  $C'=1/\pi$

$$d(I_X)_t/dX = \{[\sin^2(\bar{X}+\bar{t})]/(\bar{X}+\bar{t})^2 - (\sin^2\bar{X})/\bar{X}^2\}/\pi \dots\dots\dots (20)$$

In Figure 2, curve A represents  $(\sin^2\bar{X})/\bar{X}^2$  to an arbitrary scale. It is obvious that the representation of  $[\sin^2(\bar{X}+\bar{t})]/(\bar{X}+\bar{t})^2$  is obtained by shifting this curve in the negative direction through a distance  $\bar{X}=-\bar{t}$ . Thus curve B



represents  $[\sin^2(\bar{X} + \pi/2)]/(\bar{X} + \pi/2)^2$ , and curve C,  $[\sin^2(\bar{X} + \pi)]/(\bar{X} + \pi)^2$ . It follows from (20) that the intersections of curves A and B represent the points where stationary values of  $(I_X)_t$  for  $\bar{t} = \pi/2$  occur. Considering such of these points as lie outside the geometric image ( $\bar{X}$  positive), the first (which occurs at about  $\bar{X} = 0.85\pi$ ) gives the position of a minimum and the second that of a maximum value of  $(I_X)_t$ . When  $\bar{t} = \pi$  we see from curves A and C, Figure 2, that outside the geometric image there are no minima or maxima but only points of inflection which occur at  $\bar{X} = \pi, 2\pi, 3\pi$ , etc. From the cases, just considered, of  $\bar{t} = \pi/2$  and  $\bar{t} = \pi$ , the general conclusion may be drawn that for values of  $\bar{t}$  other than exact multiples of  $\pi$  and for points outside the geometric image there are maximum and minimum values of  $(I_X)_t$  grouped in pairs about the points  $\bar{X} = \pi, 2\pi, 3\pi$ , etc., and that the members of each pair merge into each other as  $\bar{t}$  increases. When  $\bar{t}$  is an exact multiple of  $\pi$  there are no maxima or minima outside the geometric image, but points of inflection occur at  $\bar{X} = m\pi$  where  $m$  is an integer.

(d) *Application to Resolving Power of Spectroscope*

Table 1 has a direct application to the resolving power of a spectroscope when its aperture is rectangular, e.g. when it is determined by the size of the prism. Using a slit of minimum width the usual criterion for the resolution of two spectral lines of equal intensity is that of Rayleigh, which is equivalent to the requirement that the intensity midway between the lines should not exceed about 0.81 of the central intensity of either. For this to be satisfied, the minimum separation of the lines occurs when the principal maximum of one coincides with the first minimum of the other, i.e. when  $\bar{X} = \pi$ . Let us consider what is the maximum slit width that may be used without significantly departing from this criterion.

We compute with the aid of Table 1 that for two lines separated by  $\bar{X} = \pi$  and for a slit width given by  $\bar{t} = \pi/5$  the ratio of the intensity midway between the lines to the central intensity of either is 0.823. For this slit width the central intensity of either line in the absence of the other is about one-fifth of what would be obtained with a very wide slit. With twice this width of slit the ratio is 0.86 and with a slit width corresponding to  $\bar{t} = \pi$  the ratio exceeds 1.0. The actual slit width  $\tau$  corresponding to  $\bar{t} = \pi/5$  is immediately obtained from equation (19) by substituting  $D = f$ , where  $f$  is the focal length of the collimator lens. Thus  $\tau = \lambda f/5a$ , where  $a$  is the aperture which effectively equals that of the prism. Actually a slit width considerably greater than this, viz.

$$\tau = \lambda f/a,$$

has been advocated(6) as the optimum value. It increases the light intensity to about three-quarters of the possible maximum for the sacrifice of about 20 per cent. resolving power.

From equation (19) or its equivalent here, viz.

$$\tau = \lambda \bar{f} \bar{t} / \pi a \dots\dots\dots (21)$$

it is clear that for a given relative resolving power (i.e. in relation to the maximum possible for the given prism) which is determined by the value of  $\bar{t}$ , an increase

in the aperture of the prism and hence in its absolute resolving power must be accompanied by a corresponding reduction in slit width. This does not, however, mean a reduction in the intensity despite the fact that a small slit produces, for a given prism, a small relative intensity. On the contrary, there is a gain in the absolute intensity since, as we have seen, an increase in the linear dimensions of the aperture accompanied by an inverse change in those of the source produces an increase in the intensity proportional to the change in the linear dimensions of the aperture.

In this connection it may be pointed out that the intensity at the centre of an image of the slit increases continuously with increase of slit width although, as we have seen, the central intensity may for a given slit width be a local maximum or minimum. This immediately follows from the fact that its value  $I_c$  is (see equation (18))

$$\begin{aligned} I_c &= (I_{-t/2})_t = 2I_{t/2} \\ &= 2[\text{Si}(\tilde{t}) - 2\{\sin^2(\tilde{t}/2)\}/\tilde{t}]/\pi. \end{aligned}$$

Through an error in omitting the second term in this expression, van Cittert(7) gives a graph (Figure 3 in the reference quoted) of the variation of  $I_c$  with slit width which shows maximum and minimum values of  $I_c$  instead of a gradual increase in value. His argument based on this result is accordingly invalid. This error is reproduced by Sawyer(6, Fig. 46).

### 3. DIFFRACTION BY A CIRCULAR APERTURE

#### (a) General Formula

For a circular aperture the function  $\varphi(x)$  of equations (11) and (12) takes the form(1)

$$\varphi(x) = H_1(2x)/x^2 \dots\dots\dots (22)$$

where  $H_1(x)$  is Struve's function of the first order(1), so that

$$\int_0^\infty \varphi(x) dx = \pi/2,$$

and by (11)  $C' = 1/\pi$ .

$$\text{Thus } I_X = \int_0^{\bar{X}} [\{H_1(2x)\}/\pi x^2] dx \dots\dots\dots (23)$$

$$I_\infty = 1/2 \dots\dots\dots (24)$$

Tables of  $(I_X)_\infty$ , which by (15) and (24) may be expressed

$$(I_X)_\infty = 1/2 - I_{\bar{X}},$$

have been compiled by Struve(1, p. 1,016), for a range of values of  $\bar{X}$  ( $z$  in Struve's table) from 0 to 15 in steps of 0.1. These are sufficient for computing  $(I_X)_t$  by means of (15), but for ease of reference Table 4, constructed from Struve's results, is provided here.

#### (b) Central Maxima and Minima

Consider first the intensity at the centre of the geometric image. Equation (18) applies here and Table 4 immediately gives, for any value of  $\tilde{t}$  within its

TABLE 4  
VALUES OF  $I_{\bar{X}}$  FOR A CIRCULAR APERTURE

$\bar{X}$	$I_{\bar{X}}$	$\bar{X}$	$I_{\bar{X}}$	$\bar{X}$	$I_{\bar{X}}$	$\bar{X}$	$I_{\bar{X}}$
0.0	0.0000	3.8	0.4453	7.6	0.4728	11.4	0.4822
0.1	0.0270	3.9	0.4462	7.7	0.4732	11.5	0.4824
0.2	0.0539	4.0	0.4472	7.8	0.4736	11.6	0.4826
0.3	0.0805	4.1	0.4483	7.9	0.4740	11.7	0.4828
0.4	0.1066	4.2	0.4494	8.0	0.4744	11.8	0.4830
0.5	0.1322	4.3	0.4505	8.1	0.4748	11.9	0.4832
0.6	0.1576	4.4	0.4516	8.2	0.4752	12.0	0.4833
0.7	0.1813	4.5	0.4528	8.3	0.4756	12.1	0.4835
0.8	0.2045	4.6	0.4541	8.4	0.4760	12.2	0.4836
0.9	0.2268	4.7	0.4553	8.5	0.4764	12.3	0.4837
1.0	0.2479	4.8	0.4566	8.6	0.4767	12.4	0.4838
1.1	0.2679	4.9	0.4578	8.7	0.4771	12.5	0.4839
1.2	0.2868	5.0	0.4590	8.8	0.4774	12.6	0.4840
1.3	0.3044	5.1	0.4600	8.9	0.4776	12.7	0.4841
1.4	0.3207	5.2	0.4611	9.0	0.4778	12.8	0.4842
1.5	0.3358	5.3	0.4621	9.1	0.4781	12.9	0.4843
1.6	0.3496	5.4	0.4631	9.2	0.4783	13.0	0.4844
1.7	0.3621	5.5	0.4639	9.3	0.4785	13.1	0.4845
1.8	0.3735	5.6	0.4647	9.4	0.4787	13.2	0.4846
1.9	0.3837	5.7	0.4654	9.5	0.4789	13.3	0.4847
2.0	0.3927	5.8	0.4661	9.6	0.4791	13.4	0.4848
2.1	0.4007	5.9	0.4667	9.7	0.4793	13.5	0.4849
2.2	0.4077	6.0	0.4672	9.8	0.4794	13.6	0.4850
2.3	0.4138	6.1	0.4677	9.9	0.4796	13.7	0.4851
2.4	0.4190	6.2	0.4681	10.0	0.4798	13.8	0.4852
2.5	0.4235	6.3	0.4685	10.1	0.4799	13.9	0.4853
2.6	0.4272	6.4	0.4689	10.2	0.4800	14.0	0.4854
2.7	0.4304	6.5	0.4692	10.3	0.4802	14.1	0.4856
2.8	0.4330	6.6	0.4695	10.4	0.4803	14.2	0.4857
2.9	0.4352	6.7	0.4698	10.5	0.4805	14.3	0.4858
3.0	0.4370	6.8	0.4701	10.6	0.4806	14.4	0.4859
3.1	0.4385	6.9	0.4704	10.7	0.4808	14.5	0.4860
3.2	0.4398	7.0	0.4707	10.8	0.4810	14.6	0.4861
3.3	0.4409	7.1	0.4711	10.9	0.4812	14.7	0.4862
3.4	0.4419	7.2	0.4714	11.0	0.4814	14.8	0.4863
3.5	0.4428	7.3	0.4716	11.1	0.4816	14.9	0.4864
3.6	0.4436	7.4	0.4720	11.2	0.4818	15.0	0.4865
3.7	0.4444	7.5	0.4724	11.3	0.4820		

range, the required result, viz.  $2I_{\bar{X}}$  where  $\bar{X}=\bar{t}/2$ . We have seen that for the central illumination to be a maximum or minimum the expression  $[df(x)/dx]_{x=t/2}$  should be respectively negative or positive. Here  $f(x)$  is defined by (5) and (22), i.e. by

$$f(x)=A\varphi(px)=4A\{H_1(u)\}/u^2 \dots\dots\dots (25)$$

where

$$u=2px \dots\dots\dots (26)$$



From the properties of Struve's functions it may easily be shown that

$$df(x)/dx = -8Ap[3H_1(u) - uH_0(u)]/u^3 \dots\dots\dots (27)$$

where  $H_0$  refers to the Struve function of zero order. Figure 3 is the graph over a short range of the function

$$v = 3H_1(u) - uH_0(u)$$

which will be observed to be positive for values of  $u$  between 0 and 7.3, 9.4 and 13.6, etc., and negative for values between 7.3 and 9.4, 13.6 and 15.8, etc. These numbers, therefore, give the ranges of values of  $pt$  or  $\bar{t}$  within which the illumination at the centre is a local maximum or minimum respectively.

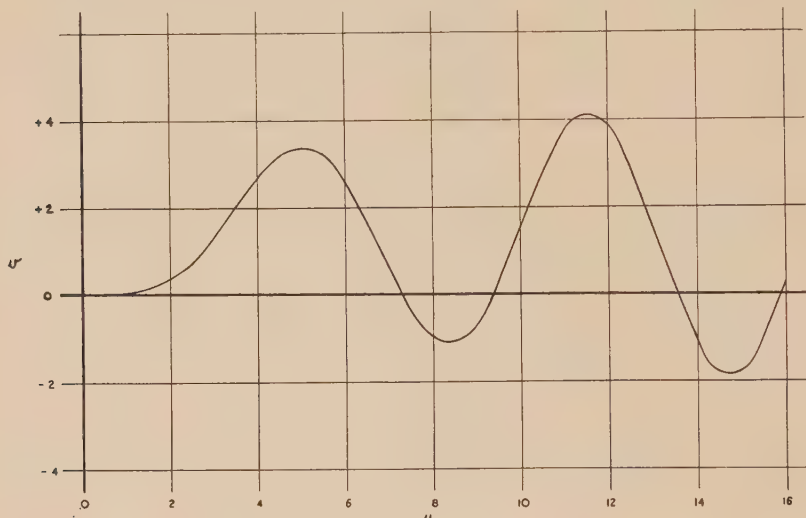


Fig. 3

From the above it is to be expected that the most favourable width of source for an observable central minimum of illumination would be about the centre of the range from 7.3 to 9.4, that is, say,  $\bar{t} = 8.4$ . Table 5 gives values of the intensity for this width for values of  $\bar{X}$  from  $-4.2$  (corresponding to the centre) to  $-2.8$  (one-sixth of the width from centre). It will be noticed that the central minimum is not sufficiently pronounced to be observable visually, as it differs from the nearby maximum at  $\bar{X} = -3.4$  by less than 0.25 per cent. The corresponding figures for the complementary case of a source, continuous except for a narrow central strip, are also shown in the table. The central maximum, which amounts to about one-tenth of the illumination for a completely continuous source, is 2 per cent. greater than the nearby minimum at  $\bar{X} = -3.4$  and is observable by eye under favourable conditions.

It is necessary to point out that the above results cannot be applied directly to the image of the cross wires or graticule lines of a telescope, since here light from a field of illumination of finite dimensions is diffracted by the cross wires or graticule lines themselves. This diffracted light must be taken into account in the final image produced. An extension of the above theory to cover these and similar cases is being undertaken.

TABLE 5  
VALUES OF  $(I_X)_t = I_{\bar{X}+\bar{t}} - I_{\bar{X}}$  FOR  $\bar{t}=8.4$  AND CORRESPONDING VALUES  
OF  $1 - (I_X)_t$ , CIRCULAR APERTURE

$\bar{X}$	$(I_X)_t$	$1 - (I_X)_t$
-4.2	0.8988	0.1012
-4.1	0.8988	0.1012
-4.0	0.8990	0.1010
-3.9	0.8994	0.1006
-3.8	0.8997	0.1003
-3.7	0.9002	0.0998
-3.6	0.9006	0.0994
-3.5	0.9009	0.0991
-3.4	0.9009	0.0991
-3.3	0.9009	0.0991
-3.2	0.9006	0.0994
-3.1	0.9001	0.0999
-3.0	0.8991	0.1009
-2.9	0.8977	0.1023
-2.8	0.8958	0.1042

#### 4. SOURCE OF SMALL WIDTH

##### (a) General Formulae

On applying Taylor's theorem to equation (7) we have for a source of small width

$$(I_X)_t = C' [\varphi(\bar{X})\bar{t} + \{d\varphi(\bar{X})/d\bar{X}\}\bar{t}^2/2 + \dots]$$

By means of this result, values of  $(I_X)_t$  may be obtained to any degree of accuracy required. If  $\bar{t}$  is so small that the term in  $\bar{t}^2$  may be neglected (22) gives, on substituting for  $C'$  the value  $1/\pi$ ,

$$(I_X)_t = \varphi(\bar{X})\bar{t}/\pi \dots\dots\dots (28)$$

Thus for a circular aperture we have from (22)

$$(I_X)_t = H_1(2\bar{X})\bar{t}/\pi\bar{X}^2 = 8\bar{t}[1/1^2.3 - (2\bar{X})^2/1^2.3^2.5 + (2\bar{X})^4/1^2.3^2.5^2.7 \dots]/\pi^2 \dots\dots\dots (29)$$

The graph of this is, of course, the well-known curve for the relative intensity at different points in the case of an infinitesimally fine line source. In addition, (29) gives the absolute intensity at any point of the diffraction pattern in terms of  $t$ , assumed small, for a source of unit brightness. When both  $\bar{X}$  and  $\bar{t}$  are small we have from (29)

$$(I_X)_t = 8\bar{t}/3\pi^2 = 0.27\bar{t} \dots\dots\dots (30)$$

As an indication of the limits over which this formula may be applied, consider a source of width corresponding to  $\bar{t}=0.2$ . In this case the illumination over the geometrical image is slightly greater than 5 per cent. of that due to a similar

continuous source. On consulting Table 4 we note that the central illumination 0.0540 agrees with the formula within the accuracy of the table. At the edge of the geometric image the figure is 0.0539. At a distance from the edge equal to the width of the source, the value is 0.0527. In view of the approximate constancy of the illumination over a distance several times the width of the so-called geometrical image, it is obvious that this term, for such small widths of source, has no physical significance.

Similar considerations apply to a rectangular aperture previously discussed. In that case, however, we have from (16)

$$\varphi(0) = \lim_{\bar{X}=0} [(\sin \bar{X})/\bar{X}]^2 = 1,$$

so that

$$(I_X)_t = \bar{t}/\pi \dots \dots \dots (31)$$

### (b) Resolution of Lines by a Telescope

Consider the resolution by a telescope of two lines of equal width and brightness. Let  $t$  be the width of the geometric images, produced by the object glass, of each line, and  $X$  the separation of the centres of the two images. The aperture here is circular, and we have for the quantity  $I_1$  which defines the illumination at the centre of each image

$$\begin{aligned} I_1 &= (I_{-t/2})_t + (I_{X-t/2})_t \\ &= 2I_{t/2} + I_{\bar{X}+\bar{t}/2} - I_{\bar{X}-\bar{t}/2} \dots \dots \dots (32) \end{aligned}$$

and for  $I_2$  which defines the illumination at a point midway between the two images

$$\begin{aligned} I_2 &= 2(I_{X/2-t/2})_t \\ &= 2(I_{\bar{X}/2+\bar{t}/2} - I_{\bar{X}/2-\bar{t}/2}) \dots \dots \dots (33) \end{aligned}$$

Applying formula (29) for small image widths

$$I_2/I_1 = 2\varphi(\bar{X}/2 - \bar{t}/2) / \{\varphi(-\bar{t}/2) + \varphi(\bar{X} - \bar{t}/2)\} \dots \dots \dots (34)$$

and in the particular case of sources of infinitely small width

$$I_2/I_1 = 2\varphi(\bar{X}/2) / \{\varphi(0) + \varphi(\bar{X})\} \dots \dots \dots (35)$$

in which  $\varphi(\bar{X})$  or  $H_1(2\bar{X})/\bar{X}^2$  is readily obtainable from tables of  $H_1(x)$ . Now when a rectangular aperture is used, as in the prism spectroscope, two infinitely fine lines are just resolved ( $I_2/I_1 = 0.81$ ) when  $\bar{X} = \pi$ . Equation (35) gives for this separation with a circular aperture

$$I_2/I_1 = 0.955,$$

so that the lines cannot be resolved. If, however, the principal maximum of one line falls on the first minimum of the other we have, approximately (see Figure 3 and equation (26))

$$\bar{X} = 3.6$$

and

$$I_2/I_1 = 0.771.$$

The actual separation for which the Rayleigh criterion of resolution  $I_2/I_1 = 0.8$  applies is given approximately by  $\bar{X} = 3.5$ .



The angular separation  $\beta$  of the lines, and their linear separation  $S$  in the object plane, are related to the quantity  $\bar{X}$  in equations (32) to (35) by an equation similar to (19), viz.

$$\beta = S/D = \lambda \bar{X} / \pi a \dots\dots\dots (36)$$

where  $D$  is the distance of the lines from the object glass and  $a$  its diameter. With this interpretation of  $a$ , (19) applies in this case to the thickness  $\tau$  of the lines. When the thickness of the lines cannot be regarded as small the ratio  $I_2/I_1$  as obtained from (32) and (33) is evaluated by means of Table 4.

As an example, consider the separation given by  $\bar{X}=3\cdot6$ . For widths of source given by  $t=0\cdot0, 0\cdot2, 0\cdot4, 0\cdot6$ , the ratio  $I_2/I_1$  then has the respective values  $0\cdot771, 0\cdot777, 0\cdot775$  and  $0\cdot781$ . It will be noticed that it is not an invariable rule that the resolution decreases with increase in the width of the lines, although owing to the small magnitude of the effect this is of purely formal interest.

#### IV. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Physics, C.S.I.R.

#### V. REFERENCES

- (1) STRUVE, H.—*Ann. Phys. Chem.* **17**: 1008 (1882).
- (2) RAYLEIGH, J. W. S.—“Scientific Papers”, Vol. 3, 1st Ed., p. 81. (Cambridge Univ. Press, 1902.)
- (3) DRUDE, P.—“The Theory of Optics”, p. 217. Translated by C. R. Mann and R. A. Millikan. (Longmans, Green & Co.: London, 1933.)
- (4) “Table of Sine and Cosine Integrals.” Prepared by the Federal Works Agency, Works Projects Administration, for the City of New York, 1942. (U.S.-N.B.S., Maths. Tables Project, M.T.13.)
- (5) SCHWERD.—“Beugungerscheinungen.” (Mannheim, 1835.)
- (6) SAWYER, R. A.—“Experimental Spectroscopy”, 1st Ed., p. 107. (Prentice-Hall Inc.: New York, 1944.)
- (7) VAN CITTERT, P. H.—*Z. Phys.* **65**: 547 (1930).

# THE ANGULAR DISTRIBUTION OF D-D NEUTRONS SCATTERED BY DEUTERONS

By J. F. DARBY\* and J. B. SWAN\*

(Plate 1)

[Accepted for Publication October 15, 1947]

## *Summary*

The angular distribution of D-D neutrons scattered by deuterons has been investigated using a Wilson cloud chamber. The apparatus and experimental procedure, and the method of selecting and reducing the observations, is described. Comparison of the result with the calculations of Buckingham and Massey(3) indicate better agreement with the distribution to be expected assuming ordinary forces than with that assuming exchange forces.

## I. INTRODUCTION

A direct and convenient method for investigating the nature of nuclear forces, in particular the neutron-proton interaction, is to observe the scattering of the elementary particles by light nuclei, the simplest case being the scattering of neutrons by protons. In such an experiment it is most convenient if all the neutrons are of the same energy, and the only convenient laboratory source of homogeneous neutrons is the deuteron-deuteron reaction, which gives neutrons of energy about 2.5 MV. If the scattering of neutrons by protons be observed, this energy is too low for the difference between exchange and ordinary forces to appear in the angular distribution of the scattered particles or the recoiling nuclei. For, at this energy, the scattering is mainly *s*-scattering—only those particles of zero relative angular momentum can approach one another sufficiently closely to interact—and theory predicts for the two types of force angular distributions which do not differ enough to be readily detected by experiment(1, 2).

If, however, the collisions are between neutrons and deuterons, a neutron of 2.5 MV energy having 1 quantised unit of angular momentum relative to the deuteron can approach closely enough to interact strongly, and *p*-scattering may be expected to occur sufficiently often to influence the distribution. In this case, theory predicts a marked difference between the angular distributions to be expected on the respective assumptions of exchange and ordinary forces(3).

Previous results(4, 5, 6), more fully described below, agree qualitatively, except for one observation giving an isotropic distribution, in finding a maximum for back-scattering of the neutron, but differ in the prominence of this maximum

\* Physics Department, University of Melbourne.

and in the detailed shape of the rest of the differential cross-section—angle of scattering curve.

## II. EXPERIMENTAL

The method used was to photograph the tracks of recoil deuterons in a Wilson cloud chamber filled with heavy methane ( $\text{CD}_4$  95%, deuterium content of hydrogen present 99.3%) and containing a few cc. of heavy water to provide the necessary vapour. The neutron source was a 200 KV D-D generator previously described(8), with the addition of a contactor in the primary circuit of the transformer supplying the ion-source voltage for pulsing the deuteron beam and hence the neutron output. This contactor was operated by the main cam shaft which controlled the operating sequence of the cloud chamber as described below.

The cloud chamber (Fig. 1) was of the rubber diaphragm type, 25 cm. in diameter and 5 cm. deep; the floor A consisted of a brass disc perforated by

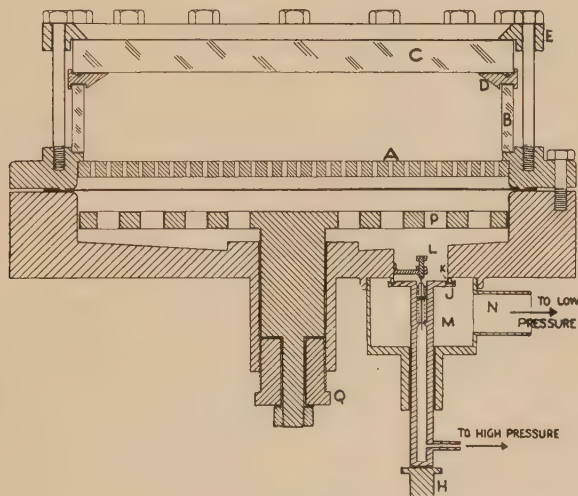


Fig. 1.—Diagram of Wilson cloud chamber.

many  $\frac{1}{16}$  in. holes covered with black velvet, the wall B was a section of a  $\frac{1}{4}$  in. thick glass cylinder, and the roof C a disc of  $\frac{1}{2}$  in. plate glass. The wall was waxed to the floor and to a seating ring D which carried a rubber gasket on which the roof was held by a brass closing ring E. The operation of the chamber was as follows. The armature of a pot magnet (not shown) pushes up the rod H to press the main valve plate J with its gasket against the seating K. The pin L then opens the Schraeder valve M to connect the space below the diaphragm with a high (constant) pressure tank (not shown), and the diaphragm is forced upward. When the magnet current is interrupted, the main valve plate is forced away from its seating and the space below the diaphragm is opened through N allowing an expansion. The expansion ratio is controlled by raising



or lowering the perforated plate P by means of the nut Q\*. A clearing field of about 200 volts was applied to a ring of aluminium foil cemented to the glass roof just inside the ring D. The whole depth of the chamber was illuminated by two heated capillary mercury arcs of the type described by Dahl, Hafstad, and Tuve(9) with cylindrical condensing lenses before them. These lights were flashed for about 1/10 second just after the expansion from a 7000-volt transformer.

The sequence of operations is controlled by electrical contacts operated by adjustable cams on a shaft driven through reduction gearing at about 1 r.p.m. by an electric motor. One set of contacts breaks the magnet current and starts an electronic delay circuit which, at a preset instant later, makes and breaks the primary circuit of the lamp transformer. A second set removes the clearing field and a third operates the ion beam contactor of the neutron generator to produce a pulse of neutrons. An intermittent chain drive from the cam shaft operates the camera mechanism, opening the shutters for the duration of the light flash and winding on the film after the flash.

Photographs were taken on Ilford HPX 35-mm. film in two Ica Kinemo cameras set with their axes at right angles to one another and to the line joining the centre of the chamber to the generator target and at 45° to the plane of the cloud chamber. The distance from target to chamber centre was 1 metre and neutrons emitted at right angles to the direction (vertical) of the deuteron beam could enter. A pair of photographs is shown in Plate 1. The neutron beam is coming from the bottom.

### III. MEASUREMENT OF TRACKS

The angle of recoil of each track relative to the direction of the incident neutron was found by reprojecting the pair of photographs on to a special screen. The reprojection was done, using the same lenses as were used in the cameras, by two projectors which were aligned in the following way. A number of concentric circles and diameters were drawn on white card, the drawing was placed in the centre of the chamber, and photographed. The drawing was then placed on the reprojection screen and the photographs in the projectors, which were then adjusted until each reprojected image of the drawing completely coincided with the drawing. It was then assumed that each image on the screen would occupy the same position relative to the projector as the track in the chamber had occupied relative to the camera. The measuring machine is shown in Figure 2a. A carriage A slides on two vees and a flat along the rails B and carries a cradle C which can rotate about an axis D perpendicular to the rails, the rotation being measured by a graduated circle E. The rail assembly can be swung in a horizontal plane and be tilted in a vertical plane by the jack F about the universal joint G, the centre of which represents the neutron source. The cradle C carries a rotatable white screen H with a radius drawn upon it. Rotations of the screen and radial line are measured by the graduated circle J.

\* We are indebted to Dr. J. C. Bower for the details of this design of chamber control which was developed at the Cavendish Laboratory.

The axis D is in the plane of H and passes through its axis of rotation. The line joining the centre of H and the centre of G is parallel to the rails and represents the direction of the incident neutron. The distance from G to the plane formed by the optic axes of the projectors was made the same as the distance from the generator target to the plane of the camera axes—1 metre.

The photographs are projected on to the screen H, which is then moved until the radial line coincides with both images of a track, whose original position and direction are then represented by the line, and the angles  $\varphi$ ,  $\theta$  in Figure 2b are read from the circles E, J, and the length of the track is measured to the nearest millimetre with a paper scale. The angle of recoil  $\sigma$  is then given by  $\cos \sigma = \cos \theta \cos \varphi$  and was read either from a nomogram or from a prepared set of tables.

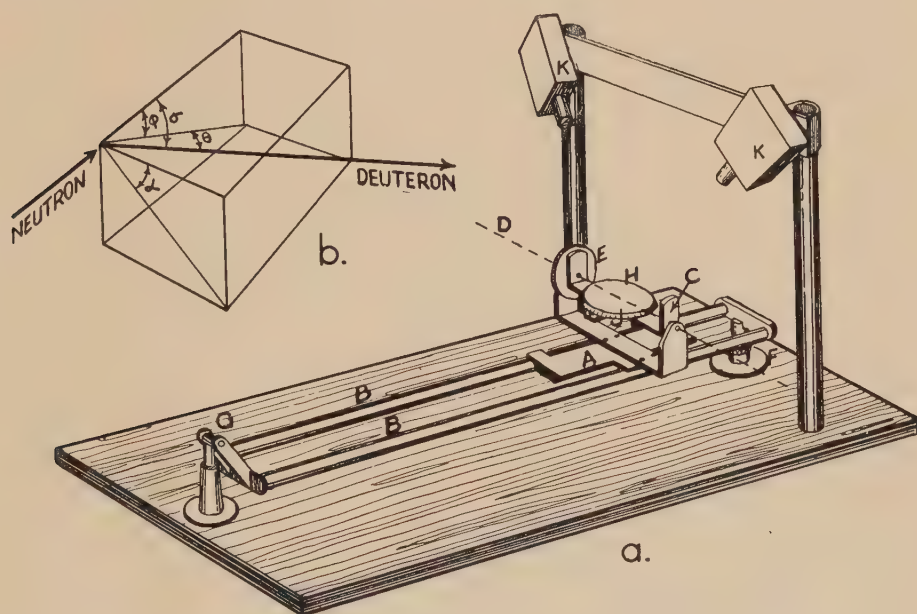


Fig. 2a.—Track measuring machine.

Fig. 2b.—Illustration of angles measured.

If the axis of rotation D were turned through a right angle so as to pass through G, then  $\sigma$  would be read directly on the circle J, while E would give the angle  $\alpha$ , but the plane H on to which the photographs are projected would then often make a small angle with the optic axis of one projector and sharp track images would not be obtained.

To avoid discriminating against long tracks which have a greater chance of hitting the walls, roof, or floor of the chamber, those tracks only were measured which started inside the shaded area in Figure 3a, from which no genuine track could reach the walls. Also, in Figure 3b, the radius vector from A to any point (say B) of the curve represents the length of a track recoiling in that direction

from a neutron travelling along OA; therefore no track was measured which started either closer than the distance  $a$  to the chamber floor and proceeded downward or closer than  $a$  to the roof and proceeded upward.

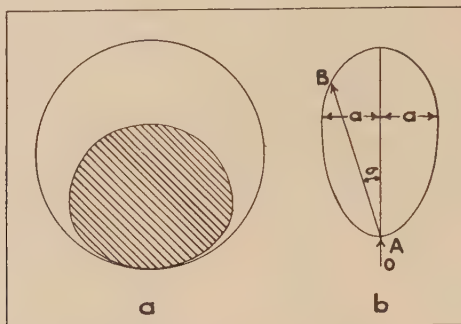


Fig. 3a.—Diagram of useful part of Wilson cloud chamber.

Fig. 3b.—Polar diagram of range of deuteron  $v$ . angle.

#### IV. SELECTION OF TRACKS

For all angles to be correct it is necessary that only those tracks be counted which are produced by recoils from neutrons which have come directly from the target. To distinguish recoils from neutrons which have been scattered into the chamber from objects in the room, the floor, walls, etc., the procedure of Dee and Gilbert(2) was followed. Each track was represented by a dot in a diagram such as Figure 4, where abscissæ represent angles of projection and the ordinates track lengths. These points should lie on a smooth curve relating angle of recoil and range. The equation  $E_d = 8/9 E_n \cos^2 \sigma$ , where  $E_d$  represents the energy of the deuteron and  $E_n$  that of the neutron; the range-energy relation for deuterons and the nature and pressure of the gas in the chamber determine this curve. In fact, allowing for a slight energy spread of the neutrons arising from the use of a thick target in which deuterons of all energies from 0 to 200 KV could cause disintegrations, for the angle subtended at the target by the chamber and for errors of measurement, the points will be concentrated about and between two such curves, one exhibiting angle  $v$  range in the unexpanded chamber and the other the same in the expanded chamber, since neutrons could enter at any instant during the expansion. The other points, distributed more or less at random over the whole diagram, are assumed to be recoils from neutrons which have been scattered into the chamber, perhaps losing energy in the process. Dee and Gilbert have shown that the numbers of particles scattered per unit solid angle in centre of gravity coordinates is obtained directly from observations in room coordinates by counting the numbers of tracks in equal intervals of  $\cos 2 \sigma$ . The numbers in such intervals between certain inclusion curves shown in Figure 4 were therefore counted to give the result. To assist in drawing these curves, a number-range histogram such as Figure 5 was constructed for each angle interval, and the inclusion curves, of the same shape as the theoretical range-angle curves, drawn so as to embrace the peaks appearing in the histograms.



A correction for recoils from scattered neutrons appearing in the allowed band was made proportional to the density of points just outside the exclusion curves, by subtracting the difference between the numbers in the first count of admitted tracks and the numbers in a band twice as wide, from the first count. In Table 1 this is called the random correction.

The earliest observations showed a much smaller concentration about the theoretical curves than does Figure 4, which was attributed to a large background of scattered neutrons in the room and was much improved by surrounding the generator target with 10 cm. of paraffin wax, leaving a small hole towards the

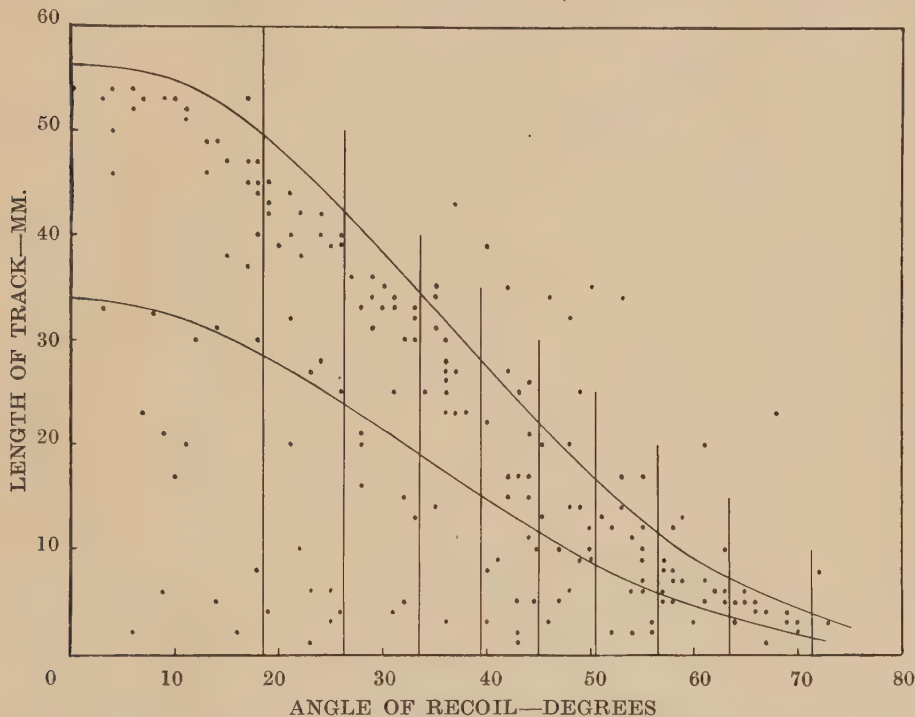


Fig. 4.—Diagram exhibiting range and angle of measured tracks, inclusion curves and equal intervals of  $\cos 2\sigma$ .

cloud chamber, and building a wall partly of 23 cm. of water and partly of 10 cm. of paraffin in front of the chamber with a slot facing the target. It is probable also that the use of fairly narrow tubes as electrodes in the neutron generator allowed an appreciable fraction of the beam to strike them, giving rise to secondary neutron sources.

A few long tracks attributed to recoil protons were observed and a correction was made at the larger recoil angles where allowed bands for protons and deuterons would overlap. The mean number of long tracks in the first three intervals was taken as the number to be expected in each interval—the distribution of protons recoiling from neutrons of this energy is known to be isotropic in centre of gravity coordinates—and the correction was obtained by multiplying this

number—10—by the ratio of the overlap area of the band to the allowed deuteron area in each interval.

#### V. RANDOMNESS OF TRACKS

A plot of points was made displaying in cartesian coordinates the  $\varphi$  and  $\theta$  values for each track. This diagram should show a uniform density of points in areas bounded by loci of constant  $\sigma$ . It showed a preponderance of tracks with  $\varphi < 0$ , i.e. tracks going down. This was finally traced to the fact that the slot in the wall before the chamber obscured the lower part of the chamber, so that the volume in which acceptable tracks going upwards could start was

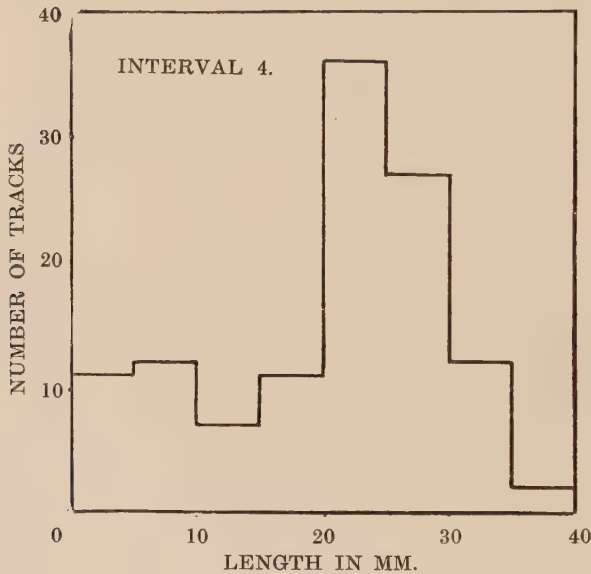


Fig. 5.—A number-range histogram.

smaller than the corresponding volume for tracks going down. Aside from this, the departure of the number of tracks per quadrant of this diagram from uniformity was small enough to have arisen from random errors of sampling ( $\chi^2$ -test), and the difference between the result for the angular distribution given by the  $\varphi < 0$  and  $\varphi > 0$  groups of tracks was similarly small, except in the last interval, No. 9,  $\cos 2\sigma = -0.6$  to  $\cos 2\sigma = -0.8$ , which must be treated with reserve in any case because of the shortness of the tracks ( $< 5$  mm.) in this interval, the consequent difficulty in measuring them, and the probability of missing some or many altogether.

#### VI. RESULTS

A preliminary experiment was carried out with ordinary methane in the cloud chamber, and the proton distribution found agreed with the isotropic distribution found by other observers(2, 10, 11). The distribution of the deuteron tracks selected by the above method from 1,503 measured is shown in

Table 1 and Figure 6. An idea of the statistical spread was obtained by calculating the standard deviation using the formula for a binomial distribution  $s = (Npq)^{\frac{1}{2}}$  where  $N$  is the total number of tracks,  $p$  is the probability that a track will lie in a particular interval, and  $q (=1-p)$  is the probability that a track will not lie in that interval. For a non-isotropic distribution,  $p$  is different for each interval, and the best estimate of  $p$  was taken to be the ratio of the number of tracks in any interval to the total number, so that the formula becomes

$$S_i = \left\{ n_i \left( 1 - \frac{n_i}{N} \right) \right\}^{\frac{1}{2}} \quad i=1, 2, 3 \dots 9$$

TABLE 1

Interval	Cos $2\sigma$	No. of Tracks	Proton Correction	Random Correction	Corrected Number	"Standard Deviation"
1	1.0-0.8	217	0	26	191	11.7
2	0.8-0.6	134	0	15	119	10.0
3	0.6-0.4	109	0	17	92	9.0
4	0.4-0.2	74	1	13	60	7.4
5	0.2-0.0	51	2	18	31	5.5
6	0.0- -0.2	72	3	14	55	7.1
7	-0.2- -0.4	74	5	24	45	6.5
8	-0.4- -0.6	103	5	27	71	8.0
9	-0.6- -0.8	68	6	9	53	7.0

For the reasons given above, the number given for interval No. 9 is at best only a lower limit. The large effect of the random correction indicates that the overall precision of the observation is probably low. The ratio between the numbers in intervals 1 and 5 is changed by it from about 4 : 1 to about 6 : 1.

The points found were fitted to the curves given by Buckingham and Massey(3) for both types of force, better agreement being found—ignoring point 9—with the curve for ordinary forces. The curves and the observed points, before and after correction, are shown in Figure 6.

## VII. COMPARISON WITH PREVIOUS OBSERVATIONS

This distribution has been investigated by Barschall and Kanner(4) and Coon and Barschall(5), using ionization chambers, and Kruger, Shoupp, Watson, and Stallmann(6) using a Wilson cloud chamber. Barschall and Kanner found a distribution isotropic in centre of gravity coordinates, but stated that their resolving power was low, so that only large deviations from isotropy varying slowly with angle could have been detected. The later measurements of Coon and Barschall on about 12,000 recoils give a sharp maximum for back scattering of the neutrons, with a differential cross-section at  $180^\circ$  scattering angle more than twice that at  $110^\circ$ . These measurements extend to a neutron scattering



angle of about  $68^\circ$ , in centre of gravity coordinates, to its original direction, corresponding to an angle of recoil of the deuteron in room coordinates of  $56^\circ$  to the direction of the incident neutron. The results of Coon and Barschall have

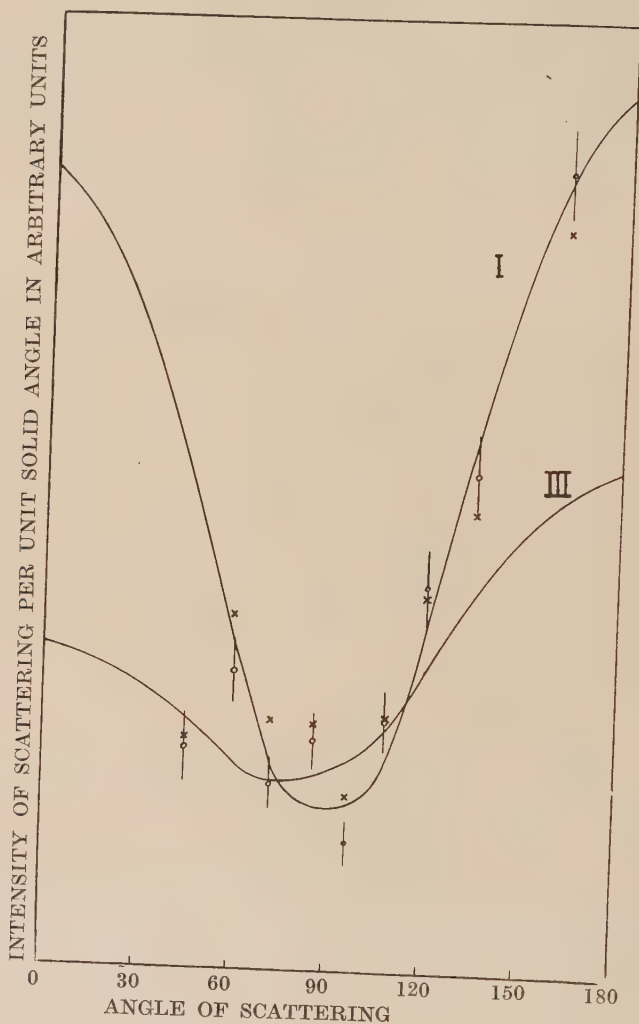


Fig. 6.—Theoretical distribution curves, I (ordinary forces) and III (exchange forces), calculated by Buckingham and Massey(3).

x=observed uncorrected points, adjusted to agree with curve I near  $120^\circ$ . o=observed points after applying corrections, adjusted to agree with I near  $120^\circ$ , and showing twice the standard deviation.

been compared with theory by Massey and Buckingham(7), who find better agreement with the distribution to be expected assuming exchange forces.

In ionization-chamber work it is necessary that all neutrons causing recoils in the chamber have the same energy. If any background of lower energy



DARBY and SWAN.—THE ANGULAR DISTRIBUTION OF D-D NEUTRONS SCATTERED BY DEUTERONS





neutrons is present, the distribution will be distorted in favour of smaller scattering angles (of the neutron in c. of g. coordinates), which would tend for neutron-deuteron collisions to reduce the relative height of the back-scattering peak.

From measurements of 328 tracks, Kruger, Shoupp, Watson, and Stallmann found a similar back-scattered maximum, but higher compared with the minimum—ratio of differential cross-sections about 4:1—which they also find near  $110^\circ$ . Their observations extend to a neutron scattering angle of  $22^\circ$  and show a second peak, about half as high as the  $180^\circ$  peak, near  $60^\circ$  in the curve of differential cross-section *v.* angle. The distribution-curve obtained does not agree well with either of the distributions calculated by Buckingham and Massey(3).

The present observations give a back-scattered maximum about six times as intense as the minimum, which appears at a slightly smaller scattering angle,  $90^\circ$ – $100^\circ$ , and agrees better with the theoretical distribution for ordinary forces.

We propose to continue the experiment with improved apparatus in the hope, particularly, of finding the distribution at small (c. of g.) angles more exactly.

#### VIII. ACKNOWLEDGMENTS

This work was suggested by Professor L. H. Martin, whom we wish to thank for advice and encouragement. We also wish to thank the State Electricity Commission of Victoria for the loan of a transformer; and the Council for Scientific and Industrial Research for a grant to enable the work to be done.

#### IX. REFERENCES

- (1) MASSEY, H. S. W., and BUCKINGHAM, R. A.—*Proc. Roy. Soc. A.* **163**: 281 (1937).
- (2) DEE, P. I., and GILBERT, C. W.—*Ibid.* **163**: 265 (1937).
- (3) BUCKINGHAM, R. A., and MASSEY, H. S. W.—*Ibid.* **179**: 123 (1941).
- (4) BARSCHALL, H. H., and KANNER, M. H.—*Phys. Rev.* **58**: 590 (1940).
- (5) COON, J. H., and BARSCHALL, H. H.—*Ibid.* **70**: 592 (1946).
- (6) KRUGER, P. G., SHOUPP, W. E., WATSON, R. E., and STALLMAN, F. W.—*Ibid.* **53**: 1014 (1938).
- (7) MASSEY, H. S. W., and BUCKINGHAM, R. A.—*Ibid.* **71**: 558 (1947).
- (8) MARTIN, L. H., HILL, R. D., and DARBY, J.—*Proc. Roy. Soc. Vict.* **58**: 135 (1947).
- (9) DAHL, O., HAFSTAD, L. R., and TUVE, M. A.—*Rev. Sci. Instrum.* **4**: 373 (1933).
- (10) KRUGER, P. G., SHOUPP, W. E., and STALLMANN, F. W.—*Phys. Rev.* **52**: 678 (1937).
- (11) BONNER, T. W.—*Ibid.* **52**: 685 (1937).

#### EXPLANATION OF PLATE 1

Stereoscopic pair of photographs of recoil tracks; the neutron beam enters at the bottom of the picture.

# THE INFLUENCE OF A HORIZONTAL WALL ON THE MOTION OF A FALLING OIL DROP

By V. D. HOPPER\* and ALISON M. GRANT\*

(Plate 1)

[Accepted for Publication January 5, 1948]

## Summary

A study has been made of the effect of a horizontal wall on the motion of oil drops of radii ranging from  $5\ \mu$  to  $15\ \mu$  falling in air. The following conclusions were reached.

Provided that  $z/a$ , the ratio of the distance  $z$  of the oil drop from the wall to its radius  $a$ , is greater than 20, the resistance  $W$  to its motion is satisfied by the formula given by Lorentz, namely :

$$W=6\pi\eta av\left(1+\frac{9}{8}\frac{a}{z}\right).$$

For  $z/a$  ranging from 2 to 500, the coefficient of  $a/z$  may be expressed empirically in the form

$$f(z/a)=k_1+k_2\frac{a}{z}$$

where  $k_1=1.08\pm0.05$  and  $k_2=1.4\pm1.0$ .

## I. INTRODUCTION

During the course of a precision determination of the electronic charge by the oil drop method it was found necessary to apply corrections for the effect of the walls on the motion of the small oil drop. Although the theory of the effect of the walls on the motion of spheres moving in a viscous medium has received considerable attention, the experiments are limited in number and the results unconvincing. Assuming the theoretical results, corrections can be made to other determinations of the electronic charge and these partly explain the discrepancies existing between them. The experimental verification of the theory is thus necessary in connection with the electronic charge determinations. It also deserves attention for its own sake in the field of hydrodynamics.

This paper deals specifically with the motion in a viscous medium of a sphere moving normally towards a plane wall. Assuming the well-known formula of Stokes(1)

$$W=6\pi\eta av\dots\dots\dots (1)$$

\* Physics Department, University of Melbourne.

† Stokes made the assumption that the inertia terms of the form  $\sigma u \frac{\partial u}{\partial x}$  ( $\sigma$  is the density of the medium) vanish in comparison with the viscosity terms of the form  $\eta \Delta u$ . Oseen(3) has

[Footnote continued on next page.]

for the resistance to a sphere moving with uniform velocity in a viscous medium, Lorentz(2) derived the following expression :

$$W=6\pi\eta av\left(1+\frac{9}{8}\frac{a}{z}\right)^* \dots\dots\dots (2)$$

for the resistance to a sphere moving normal to a plane surface,  $z$  being the distance of the sphere from the plane.

There have been only two experimental determinations of the coefficient of  $a/z$  in Lorentz's formula. Schmiedel(6) made an approximate verification by measuring the velocity over very short distances as a ball of 1 mm. diameter approached the bottom of a tank of liquid. His results indicated a much smaller correction than that predicted by the Lorentz formula. The coefficient of  $a/z$  required to explain his results would need to be 30 or 40 per cent. less than  $9/8$ .

Altrichter and Lustig(7) used Wood's metal spheres (radii  $15.6 \mu$ ,  $22.4 \mu$ ,  $34.7 \mu$ ) falling in castor oil and observed that close to the bottom of the vessel the spheres fell more slowly than would be expected from the Lorentz formula, and estimated approximately the coefficient of  $a/z$  to be  $1.267$ . They fitted a curve to their results and suggested the following empirical relation for the resistance to the motion :

$$W=6\pi\eta av\left(1+k_1\frac{a}{z}+k_2\left(\frac{a}{z}\right)^2+k_3\left(\frac{a}{z}\right)^3\right)$$

where  $k_1=1.145$ ,  $k_2=3.0$  ( $k_3$  having a negative sign and being of the order of magnitude of  $k_2$ ). Further discussion on this work will be given later.

The present investigation studied the motion of drops of radii  $5-15 \mu$  falling in air towards a horizontal glass plate. These were photographed whilst falling, using flashes of light of a duration of a few microseconds at intervals of  $1/20$ th second. The reflected images of the drops in the glass plate gave an accurate estimation of the position of the plate. Fifty drops have been studied over a range of  $z/a$  from 2 to 500. It has been found that for  $z/a > 20$  the coefficient of  $a/z$  derived by Lorentz can be considered correct (to an accuracy of the order of  $\pm 5$  per cent.). For closer distances to the plate there is a slight increase in this coefficient and an empirical formula of the form  $f(z/a)=k_1+k_2a/z$ , where  $k_1=1.08 \pm 0.05$  and  $k_2=1.4 \pm 1.0$ , has been found to fit all the data obtained.

[Continuation of footnote † from previous page.]

pointed out that without this assumption Stokes's equation must be modified to the form  $W=6\pi\eta av\left(1+\frac{3}{16}R\right)$  where  $R$  (Reynold's number) $=2av\sigma/\eta$ , the sphere being assumed to be in a medium of infinite extent. There is some doubt of the preference of this equation over that of Stokes when  $R < 1$  for spheres falling in vessels bounded by walls not too remote(4).

\* Faxén(5) has applied analysis similar to Oseen's to the case of a sphere moving along the axis of an infinite cylinder containing a viscous medium. Oseen's system of equations has not yet been applied to the case of a sphere moving normally to a plane surface in a semi-infinite medium as required for this discussion.

## II. DESCRIPTION OF EXPERIMENT

*(a) Apparatus*

The section of the apparatus in which the oil drops fell consisted of a brass cylinder, 6 cm. in diameter, which was divided into two compartments. Oil drops were produced in the upper chamber by moving a rod against the bristles of a steel wire brush, on which had been placed a small quantity of Apiezon oil B. The rod was operated by a handle outside the apparatus causing a cloud of oil drops to fall towards the bottom of the compartment where a slit 3 mm. long and 0.3 mm. wide allowed a limited number of drops to pass into the lower chamber where the actual observations were made. This slit could be closed by a magnetically operated shutter. The two chambers were clamped together and made air-tight by a rubber gasket.

In the lower chamber an optically flat, half-aluminized glass plate, 4.0 cm. in diameter, was mounted in a horizontal plane. Its position was arranged so that its top surface was 4 cm. below the slit and about 1 mm. below the axis of two cameras containing Zeiss aplanatic lenses which were mounted coaxially one on each side of the chamber. The magnification of the camera used for photographing the falling drops was not required to great accuracy. A preliminary determination had been made in connection with the electronic charge experiment and, from a study of 29 separate photographs, 15 of a Grayson ruling and 14 of quartz fibre graticules, the value 9.3816 was obtained.

The walls of the lower chamber contained three windows, one corresponding to each camera and the other admitting light to illuminate the drops. The first two were mounted perpendicular to the axis of the cameras and were made of optical glass, the surfaces being optically flat. The third window consisted of a convex lens which, together with other lenses outside the chamber, focused the light from the source on to the drops. Opposite these lenses and attached to the inner wall of the chamber was a concave mirror with a radius of curvature equal to the radius of the chamber. This reflected the light back on to the droplets and out through the window. The drops were thus illuminated almost equally in opposite directions.

*(b) Source of Illumination*

For a source of illumination two tungsten electrodes were mounted horizontally so that they were separated by a distance of 2 or 3 mm., and a trigger electrode was mounted vertically between them. A potential difference of 2000 volts across which was an 18  $\mu$ f condenser was applied to the horizontal electrodes. The triggering voltage was supplied by a multivibrator unit which was designed to step down a 1000-cycle quartz oscillator controlled signal (correct to 1 part in  $10^6$ ) to 20 pulses per second. In this way an intense source of light of short exposure time and low radiant heat content was produced at equal intervals of 1/20th second.

To avoid communicating any vibration from the spark to the apparatus the electrodes were mounted on a separate platform which rested on the floor



of the room, the table supporting the apparatus being mounted on concrete pillars which were built up from the ground and isolated from the building.

(c) *Density of the Oil and Viscosity of the Air*

The density of the oil and the viscosity of the air enter into the determination of the radius of the drop which, however, is not required to a great accuracy. (The density of the oil at 18° C. determined for the electronic charge experiment was  $0.8687_3$  g./cm.<sup>-3</sup> (temperature coefficient  $-0.00064$  g./cm.<sup>-3</sup> °C.<sup>-1</sup>). The viscosity of air has been taken as  $1.832 \times 10^{-4}$  c.g.s. units at 23.0° C. (temperature coefficient  $4.83 \times 10^{-7}$  c.g.s. units/° C.).) The air was dried by exposure to phosphorus pentoxide.

(d) *Precautions taken to maintain Constant Temperature*

Great care was taken to obtain uniformity of temperature in the apparatus as it was essential to avoid any convection currents. A 150-ohm resistance thermometer, consisting of 36-gauge cotton-covered copper wire wound around a copper tube about three inches in diameter, was placed near the apparatus. This thermometer was made one arm of a Wheatstone's bridge so that variations in temperature caused a galvanometer spot to pass over a photoelectric cell. This in turn controlled the heating system consisting of a three-kilowatt open-wire heater, which was situated several yards from the apparatus and had a fan behind it which directed the warm air against one wall of the room. Four additional fans circulated the air around the room. (One of these fans faced the observer and distributed the warm air from his body away from the apparatus.)

The method of control was a discontinuous one, the small fluctuations of temperature at the resistance thermometer being damped out by the shielding of the apparatus. The chamber containing the oil-spraying apparatus was surrounded by a layer of cotton wool which in turn was surrounded by a copper cylinder with walls  $\frac{1}{4}$  in. thick, surrounded by several layers of cloth.

In front of the window through which the light entered was a water cell in which was mounted the recently developed O.N.19 Chance radiant heat glass filter which absorbed most of the small amount of radiant heat emitted by the source.

The room (which received no direct sunlight) was darkened with blinds and all measurements were taken on dull days or at night. For the observations made at night the room was illuminated by a single sodium lamp to prevent irregular heating of the apparatus by radiant heat. (A sensitive Crookes radiometer placed near the apparatus was used to detect directional radiant heat and no movement of the vanes occurred during the progress of an experiment.)

When all these precautions were taken it was found that the maximum daily temperature fluctuation inside the apparatus, measured by a copper-constantan thermocouple connected to a Wolff Thermokraftfrei potentiometer, was of the order of  $1/200^\circ$  C. After runs of illumination of several minutes no variation in temperature was detected. Two mercury thermometers reading to

0.05° C., one above and one below the apparatus, gave identical readings indicating good distribution of external air temperature.

(e) *Experimental Procedure*

For this experiment one camera was used for photographing the oil drops, the other being used as a short-focus telescope for observation during the actual exposure. The depth of focus of the telescope was such that if a drop appeared in focus with the eye accommodated on a cross wire, the drop was also focused on the photographic plate. The image of the drop when focused consisted of two bright points corresponding to light scattered from regions near the ends of the horizontal diameter. To avoid errors due to depth of focus, measurements were made on photographs showing the two separate images. As the drop approached the horizontal plate, reflected images of the drop appeared, and from the photograph of these, the distance of the drop from the plate could be estimated. Trouble was at first encountered owing to the collection of oil on the surface of the plate immediately below the slit. Plate 1 is a photograph showing several drops of different sizes, illuminated at 20 flashes per second, falling a distance of 3 mm. towards the horizontal plate on which an excessive amount of oil had collected. The reflected images of a drop in the plane are as sharply defined as the impressions of the drop above the plane. The position of the plane estimated from these reflected images, however, will not exactly correspond to the distance of the drop from its surface owing to the collection of the small layer of oil. To avoid this small error, the substitution of an oil bath in place of the glass plate was tried, but it was found that the slightest movement of the oil surface was communicated to the air above it, causing irregular motion of the falling drops. It was decided to persevere with the glass plate and to take the following precautions to avoid the collection of sufficient oil on the plate to affect the measurements.

The slit was adjusted so that it was at a small angle to the object plane of the camera. Using a clean plate an exposure was taken until three or four falling drops in focus had passed through the field of view and then the slit was closed. Usually the total exposure time occupied one or two minutes and the chance of one drop being close enough to influence the motion of another was remote. The slit was next rotated so that a clean surface of the glass plate was immediately below the section of the slit in the object plane of the camera. After a second photograph was taken the apparatus was dismantled and the plate cleaned. The next pair of exposures was taken after a delay of at least 1½ hours when the glass had reached temperature equilibrium with its surroundings.

The photographs obtained by this method showed the presence of only a few small deposits of oil on the plate, each deposit being produced by a single oil drop. The surface of the plate immediately below a falling drop was free of oil and its position could therefore be accurately measured by means of the reflected images of the drop in the plane. These photographs were measured by means of a Hilger travelling microscope which had been carefully calibrated over the full length of the thread in terms of a certified N.P.L. standard metre.

Fifty oil drops were studied using this method, 23 having radii between  $5\ \mu$  and  $9\ \mu$ , 16 between  $9\ \mu$  and  $11\ \mu$ , and 11 between  $11\ \mu$  and  $17\ \mu$ .

(f) *Study of Emulsion Shift*

Emulsion shift produced during development and drying was measured by placing in front of the film a glass graticule consisting of parallel lines forming squares having spacings of 1 cm. The background illumination was sufficient to give a clear image of this graticule on the photograph. Measurement of the spacings of the lines on the developed negatives revealed an overall contraction of the order of 1 part in 1000. A correction of this magnitude to the magnification would be necessary, but the value 9.38 was used in the calculations, the error arising by neglecting this correction being negligible. The prints of the rulings on the films were straight lines indicating that local irregularities were small. This was true even near the more heavily exposed sections of the film, i.e. near the image of the surface of the plate where irregularities due to the "gelatine effect" and "turbidity effect" might be expected.

### III. ANALYSIS OF RESULTS

(a) *Theory of Method of Analysis*

For the analysis of the results the assumption is made that the resistance to a sphere falling in an ideal viscous medium towards a horizontal plane is of the form

$$W = 6\pi\eta av(1 + f(z/a)a/z) \dots \dots \dots (3)$$

where the coefficient of  $a/z$ ,  $9/8$  in Lorentz's equation, is replaced by  $f(z/a)$  some function of  $z/a$ , which is to be determined.

Equation (3) must be modified slightly to allow for slip and effects due to the cylindrical side wall and the top surface of the cylindrical container which encloses the air through which the oil drops fall. It is assumed that the oil drops are perfectly spherical. Equation (3) then takes the form

$$W = 6\pi\eta avS(a) \left(1 + f\left(\frac{z}{a}\right)\frac{a}{z}\right)^* \dots \dots \dots (4)$$

where  $S(a) = \left(1 + 2 \cdot 1 \frac{a}{l} + \frac{9}{8} \frac{a}{h} - \frac{b}{pa}\right)$ ;  $\left(2 \cdot 1 \frac{a}{l} + \frac{9}{8} \frac{a}{h}\right)$  is the correction for the cylindrical walls and the top surface, where  $l$  is the radius of the cylinder and  $h$  the distance of the plate from the top surface of the container;  $-\frac{b}{pa}$  is the correction for slip, where  $b = 6.32 \times 10^{-4}$  (Davies, 8) and  $p$  cm. is the pressure of the gas.

\* This equation is true if  $R \ll 1$ . Assuming Oseen's equation in place of Stokes's equation an additional factor of  $\left(1 + \frac{3}{16}R\right)$  is required on the right-hand side of equation (4). Since this correction will only produce a small variation in the estimation of the radius of the drop it will not appreciably affect the value of  $f(z/a)$ .

Allowing for the change in velocity of the drop as it reaches the plate, the equation of its motion may be written in the form

$$-\left(\frac{\rho + \frac{1}{2}\sigma}{\rho}\right)m\ddot{z} - \frac{m(\rho - \sigma)}{\rho}g = 6\pi\eta a S(a)\dot{z}\left(1 + f\left(\frac{z}{a}\right)\frac{a}{z}\right) \quad \dots\dots\dots (5)$$

where  $z$  is the distance measured from the plane,  $\sigma$  is the density of the air, and  $\rho$  is the density of the oil drop of mass  $m$ .

If  $v_0$  is the terminal velocity in the absence of the horizontal plate

$$\frac{m(\rho - \sigma)g}{\rho} = 6\pi\eta a v_0 S(a) \quad \dots\dots\dots (6)$$

and equation (5) may be written

$$\frac{v_0}{g} \left(\frac{\rho + \frac{1}{2}\sigma}{\rho - \sigma}\right)\ddot{z} = -v_0 - \dot{z}\left(1 + f\left(\frac{z}{a}\right)\frac{a}{z}\right).$$

As, in the present instance,  $\sigma$  is small compared with  $\rho$  it is practicable to write

$$\frac{v_0}{g} \ddot{z} = -v_0 - \dot{z}\left(1 + f\left(\frac{z}{a}\right)\frac{a}{z}\right) \quad \dots\dots\dots (7)$$

Suppose  $t$  is the time taken for the drop to fall from  $z_1$  to  $z_2$ , we have, by integrating (7)

$$\int_{z_2/a}^{z_1/a} \frac{a}{z} f\left(\frac{z}{a}\right) d\left(\frac{z}{a}\right) = \frac{v_0 t}{a} - \left(\frac{z_1}{a} - \frac{z_2}{a}\right) - \frac{v_0}{ag}(\dot{z}_1 - \dot{z}_2).$$

For the analysis,  $z_1/a$  for all drops has been taken as 200. If  $v_{200}$  and  $v_{z/a}$  are the magnitude of the velocities at  $z_1/a=200$  and  $z/a$  respectively, we have

$$\int_{z/a}^{200} \frac{a}{z} f\left(\frac{z}{a}\right) d\left(\frac{z}{a}\right) = \frac{v_0 t}{a} - (200 - z/a) + \frac{v_0}{ag}(v_{200} - v_{z/a}) \quad \dots\dots\dots (8)$$

$$\text{Writing } y = \frac{v_0 t}{a} - \left(200 - \frac{z}{a}\right) + \frac{v_0}{ag}(v_{200} - v_{z/a}) \quad \dots\dots\dots (9)$$

then from equation (8),

$$\frac{dy}{d(\ln z/a)} = -f(z/a)$$

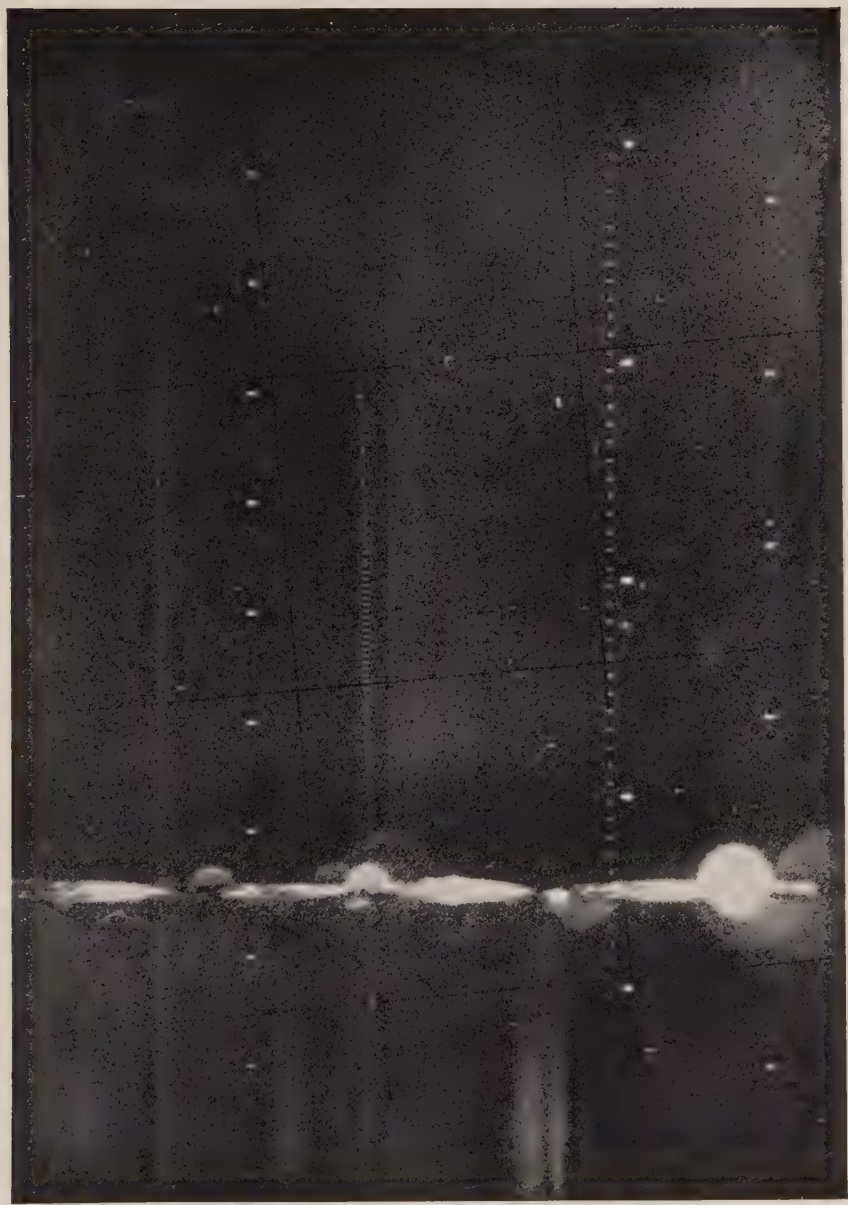
Thus if  $y$  is plotted against  $\ln z/a$ , a curve is obtained whose negative gradient gives the value of  $f(z/a)$  at that point.

#### (i) Evaluation of $y$

It will be noted that  $y$  consists of three terms :

- (1)  $\frac{v_0 t}{a}$  is the magnitude of the distance, measured in units of the radius, that the drop would traverse in time  $t$  if no horizontal plate were present.
- (2)  $(200 - z/a)$  measures the actual distance traversed by the drop from  $z_1/a=200$  in time  $t$ .
- (3)  $\frac{v_0}{ag}(v_{200} - v_{z/a})$  is a small correction term due to the inertia of the drop and we shall call this the inertia correction.





HOPPER and GRANT.—THE INFLUENCE OF A HORIZONTAL WALL ON THE MOTION OF A FALLING OIL DROP



The important terms in  $y$  are the first and second, and to estimate accurately the difference between these, which is as a maximum only 7, it is necessary to measure the time corresponding to each image of the drop photographed, the distance of each image from the horizontal plane, and  $v_0$  the velocity the drop would have if no horizontal plate were present. The time intervals are accurately known and equal  $1/20$ th second. The distance of each image from the horizontal plate can be measured precisely, since the position of the plate can be estimated from the reflection of the drop in its surface. Errors in the magnification of the camera and the estimation of the radius of the drop will not seriously affect  $y$ , because the same values will be substituted in both terms. The factor which is likely to contribute the greatest error in  $y$  is the value of  $v_0$  and this can only be derived accurately by a method of approximation.

Considering the motion of the drop at a distance from the plane such that its deceleration may be neglected, equation (7) reduces to

$$v_0 = v \left( 1 + f \left( \frac{z}{a} \right) \frac{a}{z} \right) \dots \dots \dots (10)$$

where  $v$  is the magnitude of the velocity at the position  $z/a$ . It is assumed to a first approximation that  $f(z/a)$  is  $9/8$  as given by the Lorentz theory, and  $v_0$  is calculated from the average velocity of the drop over the region  $200 < z/a < 300$  and the value of  $z/a$  corresponding to the average velocity.

The radius  $a$  of the drop is calculated from equation (6) when  $m$  is replaced by  $\frac{4\pi a^3 \rho}{3}$ . That is

$$a = \sqrt{9\eta v_0 S(a)/2g(\rho - \sigma)} \dots \dots \dots (11)$$

where  $v_0$  is its approximate value.

The inertia term presents no difficulty because only an approximate value of the difference  $v_{200} - v_{z/a}$  is required, this term being important only for large drops close to the plate.

To clarify the above method full details are given in Table 1 of the evaluation of  $y$  for each image of a drop. The same method was applied to the 50 drops studied and some 429 values of  $y$  were obtained.

To decide whether these values of  $y$  are in appreciable error owing to an incorrect choice of  $f(z/a)$  in estimating  $v_0$ , the graph of  $y$  against  $\ln z/a$  was drawn (see Figure 1). This was divided into four regions  $z/a < 10$ ;  $10 < z/a < 50$ ;  $50 < z/a < 150$ ;  $250 < z/a < 500$ . Straight lines were drawn from the point  $z_1/a = 200$ ,  $y = 0$ , to the mean of the points in the four regions, and the slope of each line  $\bar{y}/(\ln 200 - \ln z/a)$  calculated, where  $\bar{y}$  and  $\ln z/a$  are the coordinates of the mean points for each range. The values of these slopes for different sized drops and different regions are given in Table 2. Column (A) gives the slopes of the lines for the four regions using all drops, and as these are very nearly constant it can be concluded that  $f(z/a)$  must be reasonably constant over the region of  $z/a$  from 200 to a distance where the effect of the horizontal plate is negligible. The slope in the region  $250 < z/a < 500$  will only coincide with the assumed value of  $f(z/a)$  used to determine  $v_0$  from the region  $200 < z/a < 300$  if this assumed value is the true one. Column (A) indicates that there is a significant difference

TABLE 1  
CALCULATION FOR A TYPICAL OIL DROP

$t'$	$x'$	$x$	$z/a$	$\frac{v_0 t'}{a} + \frac{z}{a} - 332.63$	$\frac{v_0 t}{a} + \frac{z}{a} - 200$	$\frac{v_0}{ag}(v_{200} - v_{z/a})$	$y$
0.00	41.970	A 30.788	332.63	0.00	-0.58	0.00	-0.58
0.05	37.181	25.999	280.88	0.19	-0.39	0.00	-0.39
0.10	32.400	21.218	229.23	0.47	-0.11	0.00	-0.11
0.15	27.613	B 16.431	177.52	0.69	+0.11	0.00	+0.11
0.20	22.850	11.668	126.06	1.16	0.58	0.00	0.58
0.25	18.107	6.925	74.82	1.85	1.27	0.01	1.28
0.30	13.423	2.241	24.21	3.17	2.59	0.04	2.63
	8.941						
	4.256						

Temperature 18.4° C.; pressure 75.64 cm. Hg; magnification 9.38;

$t'$ , time in seconds measured from first position;

$x'$ , mean of three microscope readings of drop image on photograph, the last two values of  $x'$  being positions of images obtained by reflection (estimated position of horizontal plate 11.182 mm.);

$x$ , distance in mm. of drop image from horizontal plate position;

$z$ , distance of drop from plate  $\left(\frac{x}{9.38}\right)$ ;

$a$ , radius of drop ( $\sqrt{9\eta v_0 S(a)/2g(\rho - \sigma)} = 9.87 \mu$ ;

$v_0$  was determined from region A-B (see column  $x$ ) with  $k_1 = 9/8$ ;

$t$ , time in seconds measured from  $z/a = 200$ ;

$$y = \frac{v_0 t}{a} + \frac{z}{a} - 200 + \frac{v_0}{ag}(v_{200} - v_{z/a}).$$

TABLE 2

$z/a$			Magnitudes of Slopes of Lines from $z_1/a = 200$ , $y = 0$					
			$a < 9 \mu$	$9 \mu < a < 11 \mu$	$a > 11 \mu$	(A)	(B)	(C)
<10	..	..	1.16	1.19	1.19	1.17	1.15	1.15
10-50	..	..	1.13	1.20	1.14	1.15	1.12	1.12
50-150	..	..	1.10	1.22	1.12	1.14	1.10	1.10
250-500	..	..	1.18	1.21	1.09	$1.17 \pm 0.03$	$1.08 \pm 0.03$	$1.07 \pm 0.03$
$f(z/a)$ , chosen in region								
$200 < z/a < 300$			1.125	1.125	1.125	1.125	1.08	1.08

(A) Slopes for all drops in regions of  $z/a$  given, with  $f(z/a) = 1.125$ .

(B) As for (A) but with  $f(z/a) = 1.08$ .

(C) As for (B) but corrected for influence of top surface of container.



between these two values (viz.  $1.17 \pm 0.03$  and  $1.125$ ). By assuming a different value of  $f(z/a)$  for the evaluation of  $v_0$ , a different value of the slope in the region  $250 < z/a < 500$  results. By such a method of approximation it was found that if  $f(z/a) = 1.08$  in the region  $200 < z/a < 300$  is used for estimating  $v_0$ , the value of the calculated slope in the region  $250 < z/a < 500$  is also  $1.08$ . The new values of the slopes in the four regions of  $z/a$  are given in column (B) (Table 2).

In equation (4) it was assumed that the influence of the top surface of the container on the motion of a drop was a constant, and a correction of the form  $9/8 a/h$ , where  $h = 4.0$  cm., was assumed. Actually the distance of the drop

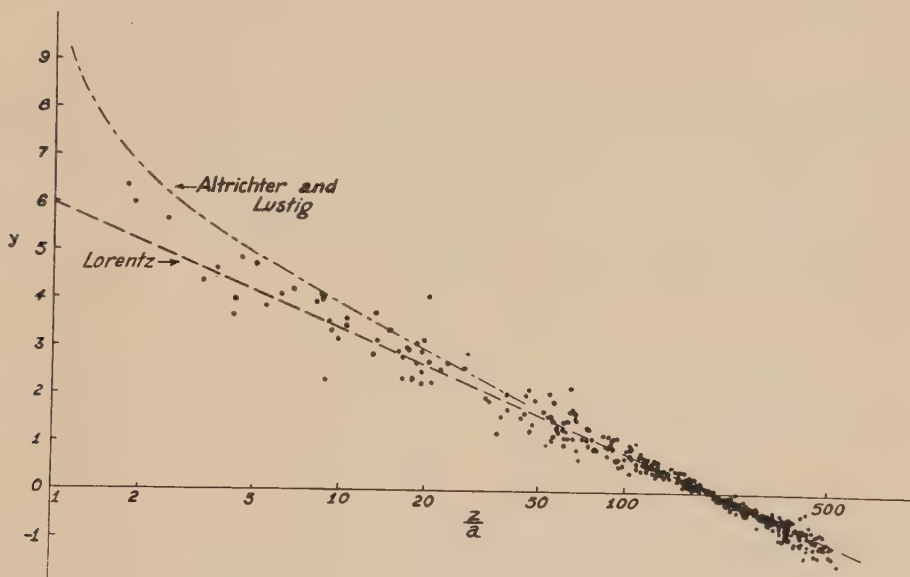


Fig. 1.—Graph of  $y$  against  $\ln z/a$  compared with theoretical result of Lorentz and the experimental result of Altrichter and Lustig.

from the surface varies from 3.6 to 4.0 cm. over the studied region of fall, and this change of distance causes a slightly varying effect on the motion of the falling drop. The values of the slopes given in column (B) are therefore influenced by the bottom horizontal plate as well as the top surface. Column (C) gives the values of the slopes corrected for this varying influence of the top surface, the only change being in the region  $250 < z/a < 500$ .

Considering column (C) it will be observed that the magnitudes of the slopes of the lines increase slightly as  $z/a$  decreases, the value of the slope in the region  $z/a < 10$  being 0.08 greater than the slope of the region  $250 < z/a < 500$ . If it is assumed that  $f(z/a) = k_1 + k_2 a/z$  it is possible to evaluate  $k_1$  and  $k_2$  by the following method.

Using equation (8)

$$y = \int_{z/a}^{200} a/z f\left(\frac{z}{a}\right) d(z/a)$$

then by substitution

$$y = k_1 \ln\left(\frac{200}{z/a}\right) + \left(\frac{a}{z} - \frac{1}{200}\right) k_2$$

or

$$\frac{y}{\ln\left(\frac{200}{z/a}\right)} = k_1 + \frac{\left(\frac{a}{z} - \frac{1}{200}\right)}{\ln\left(\frac{200}{z/a}\right)} k_2 \dots\dots\dots (12)$$

$-\frac{y}{\ln\left(\frac{200}{z/a}\right)}$  is the slope of a line drawn from  $y, z/a$ , to the point  $y=0, z/a=200$ .

Using the value of the slopes in column (C) and the corresponding mean values of  $z/a$ , four linear equations can be obtained by substituting in (12), and from these  $k_1$  and  $k_2$  can be derived by least square theory.

This analysis gives  $k_1 = 1.08 \pm 0.01$  and  $k_2 = 1.4 \pm 0.3$ . The errors do not take into account the error in estimating  $v_0$  which in turn is due to the error in the slope of the line in the region  $250 < z/a < 500$ . Allowing for this,  $f(z/a)$  may be expressed in the form

$$f\left(\frac{z}{a}\right) = k_1 + k_2 \frac{a}{z}$$

where  $k_1 = 1.08 \pm 0.05$  and  $k_2 = 1.4 \pm 1.0$ .

It will be observed that over the region  $z/a > 20$ ,  $f(z/a)$  does not differ from  $9/8$  by greater than the error  $\pm 0.05$ .

## (ii) *Alternative Method of Analysis.*

Since the value of  $f(z/a)$  has been shown to be practically constant over the region of  $z/a$  studied, a check can be made by a method which does not involve the approximation method for estimating  $v_0$ .

Modifying equation (10) by assuming  $f(z/a)$  is a constant  $k$ , gives

$$v = v_0 / (1 + k a/z).$$

Integrating with respect to  $t$ ,

$$t_1 = \frac{1}{v_0} \int_{z_1}^{z_2} (1 + k a/z) dz$$

where  $t_1$  is the time taken for a drop to fall from  $z=z_1$  to  $z=z_2$ .

That is

$$t_1 = \frac{1}{v_0} \{z_1 - z_2 + ka(\ln z_1 - \ln z_2)\}.$$

Considering three positions of the drop corresponding to  $z_1, z_2$ , and  $z_3$  where the time interval  $t$  between  $z_1$  and  $z_2$  is equal to that between  $z_2$  and  $z_3$ , then

$$\begin{aligned} t &= \frac{1}{v_0} \{(z_1 - z_2) + ka(\ln z_1 - \ln z_2)\} \\ &= \frac{1}{v_0} \{(z_2 - z_3) + ka(\ln z_2 - \ln z_3)\} \end{aligned}$$

whence

$$k = \frac{z_1 - 2z_2 + z_3}{a \ln(z_2^2/z_1 z_3)}.$$

Thus from three positions of the image of the drop on the photographic plate corresponding to equal time intervals, the value of  $k$  can be estimated.

The values of  $k$  for all drops having at least one value of  $z/a$  between 8 and 50 were calculated by this method, a small correction being made where necessary for the inertia term. The mean values obtained for the three ranges of sizes of drops are as follows:

$a > 11 \mu$	$k = 1.16$
$11 \mu > a > 9 \mu$	$k = 1.17$
$a < 9 \mu$	$k = 1.11.$

The average value of  $k$  for all drops is  $1.14 \pm 0.05$ , in good agreement with the average slope 1.12 of the curve obtained by plotting  $y$  against  $\ln(z/a)$  over the same region of  $z/a$ .

#### IV. CONCLUSION

A study has been made of the effect of a horizontal wall on the motion of oil drops of radii  $5 \mu$ - $15 \mu$  falling in air. The following conclusions were reached from this work:

1. Provided that  $z/a$ , the ratio of the distance of the oil drop from the wall to its radius, is greater than 20 the resistance to its motion obeys the formula given by Lorentz,

$$W = 6\pi\eta av \left(1 + \frac{9}{8} \frac{a}{z}\right),$$

when allowance is made for the slip term at atmospheric pressure.

2. Including values of  $z/a < 20$  it is deduced that the coefficient of  $a/z$  in the above formula may be expressed in the form

$$f(z/a) = k_1 + k_2 a/z$$

where  $k_1 = 1.08 \pm 0.05$  and  $k_2 = 1.4 \pm 1.0$ .

There has been very little experimental work carried out previously on the problem of the influence of a horizontal plate on the motion of a falling sphere. The only planned work meriting attention is that of Altrichter and Lustig who studied the motion of Wood's metal spheres falling very slowly in castor oil.

They derived an empirical relation of the form  $f(z/a) = k_1 + k_2 a/z + k_3 \left(\frac{a}{z}\right)^2$ , where  $k_1 = 1.145$ ,  $k_2 = 3.0$ , and  $k_3 \approx -3$ ; this result was obtained by assuming in their determination a value of  $f(z/a)$  of 1.125 at a distance of 5 mm. ( $z/a = 200$ ) from the plate. Substituting  $z/a = 200$  in their empirical relation gives a value of the function at this point of 1.160. Had their analysis been carried out completely, the derived value (1.160) and the assumed value (1.125) at this value of  $z/a$  would agree. The data given are insufficient to correct their work but a rough calculation shows that  $k_1$  would be decreased somewhat and  $k_2$

increased. For analysing their results they used an alternative method which is similar to the second method described in this paper. This method assumes that  $f(z/a)=k$ , a constant, over the range of  $z/a$  considered. They derived  $k=1.267$  but as no details are given of the range of  $z/a$  used in their analysis, and as their other method of analysis implies that  $f(z/a)$  is not a constant, this result is of little significance. For these reasons it is difficult to draw a definite conclusion as to their value of  $f(z/a)$  but their results show a greater variation from the Lorentz formula for small values of  $z/a$  than is obtained by the work described in this paper.

#### V. ACKNOWLEDGMENTS

The authors wish to thank Professor L. H. Martin for the keen interest he has shown in this work and for the very helpful suggestions he has made during the progress of the experiment and in the preparation of this paper.

Grants were received from the Council for Scientific and Industrial Research (Australia), and one of the authors (A.M.G.) held a Minor Research Scholarship of the University of Melbourne.

#### VI. REFERENCES

- (1) STOKES, G. G.—“Mathematical and Physical Papers”, Vols. II and III. (Cambridge Univ. Press, 1880.)
- (2) LORENTZ, H. A.—“Abhandlungen über Theoretische Physik”, Vol. I, p. 23 (1906-07).
- (3) OSEEN, C. W.—*Ark. Mat. Astr. Fys.*, Vol. 6, No. 29 (1910).
- (4) BARR, G.—“A Monograph of Viscometry”, Chap. VIII. (Oxford Univ. Press, 1931.)
- (5) FAXÉN, H.—*Ark. Mat. Astr. Fys.*, Vol. 17, No. 27 (1922).
- (6) SCHMIEDEL, J.—*Phys. Z.* **29**: 593 (1928).
- (7) ALTRICHTER, F., and LUSTIG, A.—*Ibid.* **38**: 786 (1937).
- (8) DAVIES, C. N.—*Proc. R. Phys. Soc. Lond.* **57**: 259 (1945).

#### EXPLANATION OF PLATE 1

Photograph showing images of several drops of different sizes falling towards a horizontal plane on which some oil had collected. Reflected images of the drops are seen in the lower section of the photograph. (The interval between exposures is 1/20th second and the photograph covers a distance of approximately 3 mm. above the plane.) Owing to the collection of oil on the surface of the plate these drops were not used for determining  $f(z/a)$ .

---



# ATMOSPHERIC PRESSURE CHANGES: THE IMPORTANCE OF DEVIATIONS FROM THE BALANCED (GRADIENT) WIND

By C. H. B. PRIESTLEY\*

[Accepted for Publication October 28, 1947]

## *Summary*

The paper comprises a discussion of the effect of deviations from the instantaneously balanced, or gradient, wind on the translation and development of the surface pressure field. The deviation is shown to consist primarily of two components, which are associated respectively with vertical motion and horizontal acceleration in the air trajectory.

With certain types of pressure system, the first of the above components makes a contribution towards translation of the system. But its major rôle is in maintenance of the system against the dissipating effect of surface friction. The part played by this component in the intensification of anticyclones, and the formation of secondary depressions and cold pools is discussed: in particular, these common types of development are linked with vertical motion at the tropopause and in the stratosphere. The level just below the tropopause is shown to be crucial in the study of anticyclogenesis. The collapse of an anticyclone under excessive surface heating, the formation of "heat lows", and the pressure rise associated with convective squalls are also attributed to this deviation.

By means of the deviation associated with horizontal acceleration, the changes in pressure force experienced by a moving element give rise to compensating flows which restrict any rapid local changes in surface pressure. In conjunction with the control exercised by the balanced wind(1), this accounts for the remarkable conservatism of surface pressure in face of the large convergence or divergence which atmospheric motion appears to imply.

Apart from its controlling influence, the contribution from the horizontal acceleration increases with increasing asymmetry of the system. The eastward thrust of sharp ridges of high pressure is attributable to this deviation.

## I. INTRODUCTION

The author has shown(1) that, when the streamlines of air motion are curved, the centrifugal force exercises a strong restraint on the otherwise very great tendencies towards convergence and divergence in atmospheric motion. For the purpose of studying the nature and consequences of the control so effected, it was sufficient to replace the true wind  $V$  in the pressure tendency equation by the gradient wind  $G$ , the latter being defined as that wind for which the centrifugal and deviating forces are momentarily in balance with the pressure force.

The fact that the wind can, and does, deviate from gradient does not destroy this strong controlling influence, since the deviations are rarely if ever large and

\* Section of Meteorological Physics, C.S.I.R.

systematic over a wide three-dimensional region. The linearity of the tendency equation allows the effect of the agradient component on the field of pressure change to be made the subject of a separate study. Such a study defines the scope of the present paper. Though this has been undertaken before, and though we must still be content with results mainly of a qualitative nature, it is felt that the discovery of the pressure-control inherent in the balanced component justifies a fresh examination of the importance of the unbalanced part of the wind.

We may represent these conceptions formally by writing

$$\mathbf{V} = \mathbf{G} + \mathbf{A}$$

where  $\mathbf{A}$  is the agradient or "unbalanced" component, and the equation for the local rate of change of surface pressure(2) as

$$F \frac{\partial p_0}{\partial t} = -\frac{1}{\lambda} \int_0^{p_0} \left[ G \kappa_s \frac{\partial G}{\partial s} + G^2 \frac{\partial \kappa_s}{\partial s} - \frac{2 \omega \cos \varphi}{r} V_G \right] dp - \int_0^{p_0} \frac{1}{\rho} \operatorname{div} (\rho \mathbf{A}) dp$$

.....(1)

$\rho$  represents density,  $p$  pressure and  $p_0$  surface pressure,  $G$  the magnitude and  $V_G$  the meridional component (positive if directed polewards) of  $\mathbf{G}$ ,  $r$  the distance from the centre of the earth,  $\varphi$  the latitude,  $\omega$  the earth's angular velocity of rotation, and  $\lambda = 2 \omega \sin \varphi$ .  $\kappa_s$  denotes the curvature of the streamline (positive anticyclonic, negative cyclonic),  $\frac{\partial}{\partial s}$  space differentiation along the streamline, and the integrals are taken along the vertical direction.  $F$  is a numerical factor which can be expressed in terms of the above parameters: here it suffices to quote the result that for moving pressure systems in middle latitudes it normally has a value of the order of 25. This paper will be concerned with the importance of the last term in equation (1).

## II. A DEFINITION AND A POSTULATE

The magnitude of  $F$  in equation (1) implies that there is an inherent restraint on the pressure changes which moving pressure systems can undergo. It will be shown in Section V that there is a kindred restraint on the magnitude of changes in stationary pressure systems. Accordingly the last term in (1) becomes important not when its magnitude is equivalent to a significant value of  $\frac{\partial p_0}{\partial t}$  but when it becomes comparable with terms in the first integral, which are commonly some 25 times  $\frac{\partial p_0}{\partial t}$ . This means that the method of successive approximations to  $\mathbf{A}$  can justifiably be terminated at a much earlier stage than has hitherto been recognized.

It is useful, then, to define the conception of *near-gradient wind* as applying to any wind for which the magnitude of  $\mathbf{A}$  is not greater than 0.3 times that of  $\mathbf{G}$ . In applications of near-gradient motion the squares and products of the minor (0.3) terms will then be neglected with some assurance that the main features of the field of pressure change will not thereby be obscured.

It is next postulated that *the wind is near-gradient through the main part of the atmospheric volume overlying an individual surface pressure system*. For purposes of brevity it is preferred to state this as a postulate, rather than appeal to the theoretical arguments and large amount of observational evidence which indicate its truth for the majority of systems of middle and high latitudes. Since deviations are limited in practice by the magnitude of the accelerations and vertical movements in the atmosphere (see Section III), it is to be expected that the reason for the truth of this postulate will ultimately be shown to lie in the very problem under investigation, i.e. dynamical control of the pressure distribution. The object of any present thesis must therefore lie in developing a consistent explanation of all the pressure and wind phenomena, and not in attempts to differentiate strictly between cause and effect.

The problem of pressure changes is concerned with the integrated effects of the wind in a vertical column, and for deriving the broad field of pressure change these are again integrated over a horizontal area comparable with that of the pressure system under consideration. With the object limited to a study of the broad changes, it is sufficient to examine to what extent the agradient component of the *near-gradient* wind may contribute significantly to these changes.

This limitation excludes from our present scope the consideration of purely local distortion of an established pressure pattern such as may arise from invalidity of the near-gradient assumption over a small region. The broad pressure changes may be regarded as of two types, those which are equivalent to a bodily translation of part of the pressure field, and those representing systematic development of a pressure pattern (*in situ* or in motion). This distinction, and that between systematic development and local distortion may appear somewhat arbitrary; but the latter distinction is akin to the normal statistical procedure of distinguishing between smoothed and unsmoothed data, while both are procedures which the practising meteorologist must follow in every analysis and forecast.

### III. THE EQUATION OF FRICTIONLESS MOTION

To develop an expression representing  $\mathbf{V}$  as the vector sum of  $\mathbf{G}$  and the composite deviation  $\mathbf{A}$ , we first write the equation of frictionless motion as

$$\dot{\mathbf{V}} + 2\boldsymbol{\Omega} \times \mathbf{V} = -\frac{1}{\rho} \text{grad } p + \mathbf{X} \dots\dots\dots (2)$$

where  $\dot{\mathbf{V}}$  is the acceleration,  $\boldsymbol{\Omega}$  the vector angular rotation of the earth, and  $\mathbf{X}$  the external force on the air. In the absence of friction,  $\mathbf{X}$  is a vertical force, and may be eliminated by vectorial multiplication by the unit vertical vector  $\mathbf{k}$ . We then obtain, for the southern hemisphere

$$(\mathbf{k} \times \dot{\mathbf{V}}) + 2w\boldsymbol{\Omega} + \lambda \mathbf{V} = \lambda \mathbf{J}$$

where  $\mathbf{J}$  is the geostrophic wind and  $w$  the upwards vertical velocity, and by transposition

$$\mathbf{V} = \mathbf{J} - \frac{1}{\lambda} (\mathbf{k} \times \dot{\mathbf{V}}) - \frac{2w}{\lambda} \boldsymbol{\Omega} \dots\dots\dots (3)$$

Let  $V_h$  denote the horizontal wind speed and  $\mathbf{t}$ ,  $\mathbf{n}$ ,  $\mathbf{j}$  unit horizontal vectors directed, in the southern hemisphere, respectively along the projection of the path, at right angles to the right of the path, and southwards, then

$$\mathbf{k} \times \dot{\mathbf{V}} = \mathbf{k} \times \left\{ \dot{V}_h \mathbf{t} - \kappa V_h^2 \mathbf{n} + w \frac{\partial \mathbf{V}}{\partial z} \right\}$$

and equation (3) becomes

$$\mathbf{V} = \mathbf{J} + \frac{\kappa V_h^2}{\lambda} \mathbf{t} + \frac{\dot{V}_h}{\lambda} \mathbf{n} - \frac{w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{V}}{\partial z} \right) + w \cot \varphi \mathbf{j} + w \mathbf{k} \dots \dots \dots (4)$$

$\kappa$  is used to denote the (anticyclonic) curvature of the path, and  $\dot{V}_h$  conventionally the total derivative of  $V_h$  excluding the vertically-advected acceleration component  $w \frac{\partial V_h}{\partial z}$ .

Equation (4), though of little use in itself, is exact and is in a convenient form for applying the approximation of near-gradient flow. For this we shall neglect squares and products of terms whose magnitude is not more than  $0.3G$  or, what amounts to the same thing,  $0.3J$ . The magnitudes of the various parameters which correspond to deviations of  $0.3J$  in the respective terms in equation (4) are, in latitude  $45^\circ$ :

$$\frac{\kappa V_h^2}{\lambda}: \kappa \approx \frac{1}{10 V_h} \text{ miles where } V_h \text{ is in knots.}$$

$$\frac{\dot{V}_h}{\lambda}: \text{Local rate of change } \frac{2}{3}J \text{ per 6 hours, or space change 10 knots per 100 miles.}$$

$$\frac{w}{\lambda} \frac{\partial \mathbf{V}}{\partial z}: \text{Vertical wind shear of } J \text{ per 1000 feet for } w \approx 1 \text{ cm./sec., or } \frac{J}{10} \text{ per 1000 feet for } w \approx 10 \text{ cm./sec.}$$

When viewed in the light of experience these magnitudes go far to justify the postulate of the last section.

It was shown elsewhere(1) that vertical motion can have no direct important effect on the surface pressure change, and the last term in (4) will not be further considered. The indirect effects of vertical motion on the horizontal component are represented by the two immediately preceding terms; the ratio of the second of these to the first is

$$\frac{2 \omega \cos \varphi}{\left| \frac{\partial \mathbf{V}}{\partial z} \right|}$$

which is equal to  $0.06$  in latitude  $45^\circ$  with a wind shear of  $1$  knot per  $1000$  feet. The penultimate term of (4) will accordingly be also disregarded. With these simplifications the equation becomes

$$\mathbf{V} = \mathbf{J} + \frac{\kappa V_h^2}{\lambda} \mathbf{t} + \frac{\dot{V}_h}{\lambda} \mathbf{n} - \frac{w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{V}}{\partial z} \right) \dots \dots \dots (5)$$

This differs from Sutcliffe's(3) treatment of the deviations, in which the two middle terms above are further separated by distinguishing between instantaneous



space and local time variations. This distinction is of doubtful advantage, for another term is added and the physical origin of the respective terms is somewhat obscured.

For near-gradient, or for similarly defined near-geostrophic, flow we may now write (5) in the two alternative forms

$$\mathbf{V} = \mathbf{J} + \frac{\kappa J^2}{\lambda} \mathbf{t}_i + \frac{\dot{J}}{\lambda} \mathbf{n}_i - \frac{w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{J}}{\partial z} \right) \dots\dots\dots (6)$$

$$\mathbf{V} = \mathbf{G} + \frac{\dot{G}}{\lambda} \mathbf{n}_i - \frac{w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{G}}{\partial z} \right) \dots\dots\dots (7)$$

the suffix  $i$  denoting tangent or normal to the isobar. Equation (7) represents a somewhat closer approximation than (6), but only in special circumstances is it an approximation of higher order in the mathematical sense.  $\dot{G}$  and  $\dot{J}$  are here used conventionally in the same sense that  $\dot{V}_h$  was defined, i.e. the vertically-advected acceleration is excluded.

Equation (7) is in the required form and allows the physical meaning and origin of the deviations to be readily appreciated. The first,  $\frac{\dot{G}}{\lambda} \mathbf{n}_i$ , is associated with horizontal acceleration in the trajectory. If an individual element of air experiences a change in its wind speed, it must do work or have work done upon it by a motion towards or away from higher pressure, according as it is retarded or accelerated respectively. If we conceive an "effective force" per unit mass as equal to the acceleration of an element whose motion were always that of the gradient wind, an actual element will move to the right of the additional effective force imposed upon it.

The term  $-\frac{w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{G}}{\partial z} \right)$ , or vertical derivative, applies the same concept to environmental changes experienced as a result of vertical movement. Here again the element will move at right angles to the right of the added effective force. The magnitude of this component will be proportional to the vertical velocity and to the shear of the gradient wind.

For many applications this last term is equivalent to the corresponding term in (6). As shown by Durst and Sutcliffe(4), this is equivalent to a horizontal component from the warm towards the colder environment in descending, and from cold towards warm in ascending air. It is remarkable that this elegant rule, which applies without change of wording to either hemisphere, has received so little attention.

#### IV. THE VERTICAL DERIVATIVE

##### (a) *Strict Limits of Applicability*

The derivation of the expression  $-\frac{w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{G}}{\partial z} \right)$  for the deviation which accompanies vertical motion has involved a twofold approximation. With 0.3

adopted as the ratio whose squares and products can usefully be disregarded, the limits of applicability are

$$\cos \varphi < \frac{0.3}{2\omega} \left| \frac{\partial G}{\partial z} \right| < 0.09 \frac{G}{|w|} \sin \varphi.$$

For a gradient wind shear of 1 knot per kilometre,  $\frac{0.3}{2\omega} \left| \frac{\partial G}{\partial z} \right|$  is about 1. It is, then, the second of the two inequalities which will in practice provide the more severe limitation. In cases of subsidence, save just at the centre of an anticyclone,  $\frac{G}{|w|}$  is normally 100 or more. For upslope motion along a frontal surface,  $\frac{G}{|w|}$  is not less than the inverse of the slope of the front and so will range upwards from about 50. Moreover, for fronts in equilibrium this lower value of 50 is associated with the more feeble frontal temperature contrasts and thus with the smaller values of wind shear.

The approximations are then valid in middle and high latitudes in both these types of situation for wind shears ranging from less than 1 up to 10 knots per kilometre or more, save for the rare exception of a steep front with strong vertical wind shear.

### (b) *Translation of Pressure Systems*

An exhaustive analysis of the contribution of the vertical derivative to the surface pressure change, as represented by

$$\frac{1}{F} \int_0^{p_0} \frac{1}{\rho} \operatorname{div} \left\{ \frac{\rho w}{\lambda} \left( \mathbf{k} \times \frac{\partial \mathbf{V}}{\partial z} \right) \right\} dp \dots\dots\dots (8)$$

demands an essentially synoptic approach, and the full effect on the pressure field cannot be gauged in the absence of detailed knowledge of the vertical currents. This paper is concerned with the general mechanism of pressure control, and is confined to deduction of general features of importance.

So far as the longitudinal translation of pressure systems is concerned, there is no contribution from (8) when the structure of the system and the vertical motion are symmetrical about a north-south axis. To the extent that this is so, the rules of translation derived from the principle of curvature control(1) remain unchanged.

When there is an excess of vertical motion, or an excess of wind shear, on one side of the centre the vertical derivative will in general have a corresponding asymmetry with a resulting contribution towards the translation of the system. Examples follow.

In the warm sector depression typical of middle latitudes of the northern hemisphere the vertical motion is mainly upwards and tends to be greatest in the regions where the vertical wind shear is also greatest, i.e. in the frontal zones. The vertical derivative then gives a component of motion outwards from the warm sector, at right angles to the direction of the fronts. The effect of this on development will be discussed later. So far as longitudinal translation is

concerned, if the product of rising velocity and wind shear at the cold front exceeds that at the warm, divergence would be stronger west than east of the centre, contributing towards a "retrograde" motion of the system. A greater product at the warm than at the cold front would contribute to an eastward progression.

The meridional front(5) with its associated trough typical of the Australian and New Zealand region, is in many respects a more simple system. With the warmer air situated in the northerly current to the east of the front, general ascent in the region of the trough will be accompanied by an unbalanced component from east to west. Such a transfer of air would be accompanied by falling pressure to the east and rising pressure to the west and a general translation of the trough eastwards. Since both the upward motion and the vertical wind shear, or temperature contrast, increase with increasing intensity of the trough, the trough will tend to move more rapidly the greater its intensity.

It is doubtful to what extent the regular eastward progression of anticyclones in Australian regions is an immediate dynamical effect. It is not unreasonable to regard it as purely apparent, and due to the progression of dynamically active troughs through a zone of general high pressure. Motion of the anticyclones would then appear as a dependent and residual effect. However, in anticyclones it is known that subsidence tends to be more marked to the east than to the west of the centre. In the case of warm-cored anticyclones, application of the rule developed for the vertical derivative shows that here again the unbalanced component will contribute towards an eastward translation of the high pressure centre.

### (c) *Maintenance against Surface Friction*

The vertical derivative plays an important part in the development of pressure systems, but this cannot be discussed satisfactorily unless the problem of maintenance of pressure contrasts in steady conditions be faced at the same time.

The total flow across the isobars from high towards low pressure in the friction layer is equal to(6)

$$\frac{1}{2}\rho G \sqrt{\frac{K}{\omega \sin \phi}} \sin 2\alpha = \frac{\rho G}{2B} \sin 2\alpha \dots\dots\dots (9)$$

per unit length of isobar, where  $K$  is the kinematical eddy-viscosity, assumed constant with height, and  $\alpha$  is the angle between the surface wind and isobar. The expression (9) represents a very large surface flow of air. If uncompensated by flow at another level, most pressure systems would be but transient phenomena. As an estimate of the quantities involved we may put  $K = 10^5$  cm.<sup>2</sup>/sec. and  $\alpha = 22\frac{1}{2}^\circ$ . For a gradient wind of 10 m./sec. surrounding an anticyclonic area of radius 500 km. centred in latitude  $45^\circ$  there would then be an average rate of decline of surface pressure of 3 millibars per hour within the area. Under such conditions any anticyclone of normal intensity would be destroyed in a few hours.

When the steady state is approximately maintained over a period of days, or even of hours, there must be an equal compensating flow at higher levels. In middle latitudes, in view of the positive correlation between surface pressure and tropospheric temperature, this generally involves a steady component of motion from regions of low towards regions of high pressure. No essential characteristics of the problem will be lost if from this point onwards the argument be restricted to the case of the stationary warm anticyclone. The unbalanced components of flow in such a system are shown schematically in Figure 1.

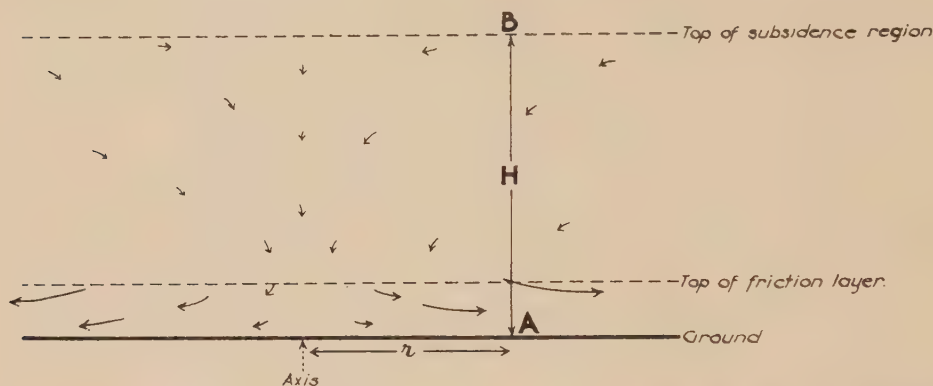


Fig. 1.—Vertical cross-section through a warm anticyclone, showing schematically the vertical, frictional, and vertical derivative components of flow. The gradient wind speed is  $G_0$  at A and  $G_0 + [G]$  at B.

The rule for the vertical derivative shows that a flow in the right sense would accompany the subsidence in such an anticyclone. To examine whether or not complete compensation can thereby be effected, we may infer from considerations of continuity that the rate of subsidence  $-w_0$  at the top of the friction layer, averaged over the circular area of radius  $r$ , is given by

$$-\pi r^2 w_0 \rho_0 = 2\pi r \frac{\rho_0 G_0}{2B} \sin 2\alpha,$$

which with the figures adopted above gives

$$-w_0 = 0.6 \text{ cm./sec.}$$

Other things being equal,  $w_0$  is proportional to  $\frac{G_0}{r}$ . Suppose the subsidence to extend from near the surface to a height  $H$  and let  $[G]$  denote the increase in  $G$  through this layer. Equating the inflow and outflow through the sides of a vertical cylinder of radius  $r$ ,

$$\begin{aligned} 2\pi r \frac{\rho_0 G_0}{2B} \sin 2\alpha &= -2\pi r \int_0^H \frac{\rho w}{\lambda} \frac{\partial G}{\partial z} dz \\ &= -\frac{2\pi r}{\lambda} \rho \bar{w} [G] \end{aligned}$$



where  $\overline{\rho w}$  denotes a mean over the area and through the column  $H$ . Then

$$\begin{aligned}\frac{[G]}{G_0} &= -\frac{\rho_0 \lambda \sin 2\alpha}{2B \overline{\rho w}} \\ &= -\frac{\rho_0 \sin 2\alpha \sqrt{K\omega \sin \varphi}}{\overline{\rho w}} \dots\dots\dots (10)\end{aligned}$$

We may now appeal to common experience. The increase of gradient wind through the subsidence layer some 10 or 20 thousand feet deep is normally about as great or rather greater than the surface gradient wind itself, i.e. the gradient wind may more than double its value between bottom and top. Setting

$$[G] \approx 1.5G_0$$

with other values as before, complete compensation will require

$$\frac{\overline{\rho w}}{\rho_0} \approx -1.1 \text{ cm./sec.}$$

Subsidence velocities of this magnitude are known to occur, and at first sight it appears that the complete internal balance is now explained on a purely dynamic basis. There is some uncertainty as to the exact value of both  $K$  and  $w$  but this in itself is no serious shortcoming, for in applying the equation of continuity to the friction layer it was evident from the form of (9) that the value of  $w$  immediately above this layer must vary as  $\sqrt{K}$ , so that in (10) any variations in the magnitudes of these quantities are mutually compensating. But the value of  $w$  required to maintain continuity in the friction layer (0.6 cm./sec.) is smaller than the 1.1 cm./sec. or higher value required above, and this discrepancy, though slight, is fatal to the completeness of the balance. There would be no such discrepancy if we set, say,

$$[G] \approx 3G_0,$$

but such a wind shear is very exceptional in the warm anticyclones of middle latitudes.

The conclusion in this case is that the vertical derivative in the region of subsidence, while of the right order of magnitude, is not adequate to compensate fully for the frictional outflow. Though not separating the vertical derivative, Durst(7) has treated a closely allied problem and his quantitative discussion indicates a similar conclusion.

#### (d) *The Importance of Radiation*

Apart from the free-atmosphere inflow associated with the subsidence which accompanies frictional outflow, there can exist further cross-isobar flows, even in steady conditions. This fact is not brought out by the simple mechanical equation of motion (2), for the latter does not take account of thermodynamic causes and effects. Assessment of the potential importance of this further component will therefore involve a short digression from the main sequence of the paper.

From considerations of the full equation of conservation of energy (6, p. 289), energy added to the air as heat ( $Q$ ) must be used in increasing the internal ( $I$ ), kinetic ( $K_e$ ), and potential ( $P$ ) energies, in overcoming the frictional force ( $\mathbf{F}$ ),

and in expansion of the volume occupied by the fluid against the pressure of its environment. In symbols,

$$J \int \rho \frac{dQ}{dt} d\tau = \frac{\partial}{\partial t} (I + K_e + P) - \int \rho \mathbf{V} \cdot \mathbf{F} d\tau + \int p V_n dS,$$

where  $J$  is the mechanical equivalent of heat,  $d\tau$  and  $dS$  the elements of volume and bounding surface to which the principle is applied, and  $V_n$  the velocity component along the outward normal to the surface. In the steady state, neglecting friction,

$$J \int \rho \frac{dQ}{dt} d\tau = \int p V_n dS,$$

an equation which may be regarded as indicating the upper limit of  $V_n$  accompanying the addition or removal of heat. In our present example we apply this equation to a circular cylinder of radius  $r$  concentric with the anticyclone.

As described by German meteorologists(8), the main non-adiabatic effect in the anticyclone is the net loss of heat by radiation from the free atmosphere.  $V_n$  is then negative, i.e. the flow is directed towards the axis. Elsasser(9), from the best data available, estimates the average radiative loss of heat from the free atmosphere as equivalent to a cooling of  $2^\circ \text{C}$ . per day. The greatest amount of inflow will occur when there is no compensating gain of heat by other means, in which case we derive from the above a value of 8 cm./sec. for  $V_n$  at 500 km. from the centre. The vertical derivative (see Sub-section IV (c)) is normally considerably greater than this.

If the "radiation-inflow" were sufficient by itself to maintain the pressure balance of the system, we would have

$$\begin{aligned} 2\pi r \frac{\rho_0 G_0}{2B} \sin 2\alpha &= \int_{z=0}^{z=\infty} \rho V_n dS \\ &= \frac{1}{RT} \int_{z=0}^{z=\infty} p V_n dS \\ &= \frac{1}{RT} \pi r^2 J C_p \int_0^\infty \rho \times (2^\circ \text{C. per day}) \times dz \dots (11) \end{aligned}$$

The radius  $r$  does not cancel in this equation. In general, then, overall balance between inflow and outflow can be achieved in this way only at a certain distance from the centre; substituting in (11) the values for  $G_0$  (10 m./sec.),  $K$ , and  $\alpha$  already adopted, this distance is estimated at 1200 km. Since this figure is arrived at from upper limiting values for  $V_n$ , it represents a lower estimate for the radius corresponding to overall balance.

#### (e) *General Aspects of Intensification and Decline*

What emerges from the two preceding sub-sections is that the radiation adjustment and the vertical derivative are both important factors in maintaining the flow necessary to compensate for surface friction. Neither in itself appears

capable, on the best evidence available, of maintaining a steady surface pressure throughout the system. They should be regarded as necessary complements of one another if the steady state is to be sustained. Any exceptionally strong manifestation of one or the other will appear as an overweight.

There is thus a fundamental basis for regarding strong subsidence in warm air as a pressure-increasing influence. Anticyclogenesis is also to be expected when radiation cooling in the free atmosphere is exceptionally pronounced, though the manner in which the scale-factor  $r$  occurs in equation (11) suggests that appreciable rises of pressure will normally result only when the strong cooling is unchecked over a wide region, e.g. the Siberian winter and Antarctic anticyclones. In any application of these principles to practical examples of anticyclogenesis, the possibility of the restriction of frictional outflow by ground contours, or by surface stabilization as represented by a decrease in the value of  $K$ , would also have to be considered.

Apart from anticyclogenesis, other applications of the dynamic rule for the vertical derivative may be made just as readily. In a warm sector depression with a predominance of ascending motion at both the warm and cold fronts, there will then be horizontal divergence away from the centre in the free atmosphere. As compared with the anticyclone, both  $w$  and the vertical wind shear are large in this case, and the deepening influence will persist so long as the occlusion of warm air continues.

In a cold depression, on the other hand, so long as convergence near the surface imposes continuation of the ascent of air, the vertical derivative will reinforce surface friction in supplying convergence which operates to weaken the surface depression. A secondary surface depression or trough at the top of an extruded warm sector will intensify, and there results the phenomenon of growth of the secondary apparently at the expense of the primary depression. As the primary fills up, the frictional convergence is weakened. This in turn reduces the ascending velocity and thereby the convergent deviation at higher levels. The stage may be reached when the depression at the surface can no longer be identified but the air column will still be cold, so that a strong cyclonic circulation persists aloft with no further tendency to fill up save by the much slower process of thermal adjustment and dissipation of its own kinetic energy. The resulting cold pools are persistent features of upper air charts.

#### *(f) Motion of the Tropopause and in the Stratosphere*

The foregoing sub-sections have dealt with the significance of the vertical derivative in the troposphere. An interesting and important application of the rule may also be made to the case of vertical motion at the tropopause and in the stratosphere.

For the warm anticyclone, already considered in some detail, there will be at higher levels a reversal of the horizontal temperature gradient, with cold stratospheric air overlying the warm central axis of the system and warmer stratosphere air at the environs of the anticyclone. This implies a reversal in sign of the vertical wind shear, and thus of the vertical derivative unless the

sense of vertical motion is also reversed. Ascent at these levels will involve horizontal convergence, tending towards an increase of the central pressure: subsidence at these levels would aid in dissipating the anticyclone. There is here a clear and purely dynamic link between the intensification or decline of a warm anticyclone and, respectively, the rise or fall of the tropopause which is so often observed as an accompaniment. From the vertical derivative, therefore, we may derive a schematic cross-section of the field of motion which is most favourable for the rapid intensification of such an anticyclone. This is shown diagrammatically with greatly magnified vertical scale in Figure 2.

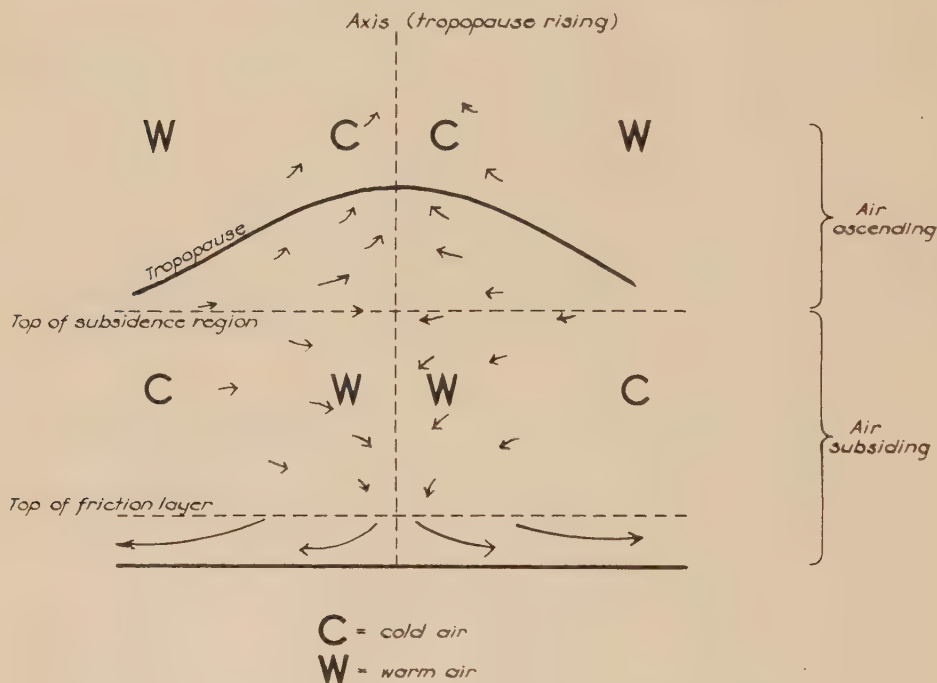


Fig. 2.—Vertical cross-section illustrating rapid anticyclogenesis.

Convergence initiated in an isolated layer in a still atmosphere would both depress the underlying air and elevate the overlying air. When convergence is deep or all-pervading there will therefore be a lower region of subsidence and an upper region of ascent. Since, as shown, this vertical motion will in turn assist in maintaining the convergence, we observe a cycle which will to a large extent be self-sustaining. In special circumstances, i.e. when the level where there is no vertical motion coincides with that at which the horizontal temperature gradient is reversed, the vertical motion and the convergence will be mutually assisting everywhere. It is reasonable to interpret this as the condition favourable for the most rapid anticyclogenesis. Empirical support for this contention may be found in the example studied by Thomas(10): the very large rates of local cooling (up to 6° C. per day) which he observed between 8 and 15 km. are



most readily attributable to a general ascent in those levels. The case analysed by van Miegham(11) shows similar characteristics, and the tropopause charts appearing in the Daily Weather Report of the British Meteorological Office provide many examples as further evidence.\* This theory and these facts point to the level just below the tropopause as one to which special attention should be directed if the intensification of anticyclones is to be better understood and forecast.

In the case of the cyclone a similar application can be made. As the secondary at the tip of the warm sector deepens at the expense of the primary depression, a lowering of the tropopause over the latter will accompany, and assist in, its decay. It is normal in such cases to observe a pronounced minimum in the height of the tropopause over the old dying primary depression or cold pool. Here again many examples are to be found in the Daily Weather Report.

#### *(g) Some Other Effects*

The strict limits of applicability of equation (6) or (7) were discussed in Sub-section IV (a), but it may be seen from equation (4) that the more general qualitative statement that rising air involves a horizontal deviation in the sense from warm towards colder regions remains valid so long as the main motion is even approximately guided by the isobars.

If the subsidence through the warm region of an anticyclone be for some reason overcome, a reversal in the sense of the vertical motion would be accompanied by a change from convergence to divergence through the depth of the now ascending air. This would then reinforce the frictional outflow. Such appears to be the mechanism which brings about the rapid disintegration of summer high pressure systems under the influence of excessive local surface heating. In the same way, excessive surface heating is responsible for the formation of "heat lows".

The temporary pressure rise associated with passing convectonal squalls has been attributed by various writers(12) to vertical acceleration, and alternatively to inflow into a region of falling pressure at the top of an air column which is contracting as a result of cooling. As neither explanation is entirely satisfactory, a third will be advanced here which depends upon ascending motion but requires neither acceleration nor a hypothetical upper isallobaric low. When the air in the central column of the squall is cooled by heavy rain, the vertical derivative indicates that continuing ascent will draw air into the column at a rate which is proportional to the velocity of ascent and which increases with the horizontal temperature gradient established by the cooling. The resulting pressure rise would then be greater in dry than in moist air on account of the

\* Since this paper went to press, Fleagle (*Journal of Meteorology*, Dec. 1947) has published a complete empirical three-dimensional study of a single anticyclone in North America. Discussing the process of anticyclogenesis, the relevant section concludes: "... cold air appeared in the lower stratosphere as a result of upward motion above the sea-level anticyclone, and simultaneously the air in the troposphere became warmer as a result of subsidence and non-adiabatic heating."

greater evaporational cooling in the former case, and would be greater in low than in high latitudes since for a given temperature gradient the deviation shown by equation (4) is roughly proportional to  $\text{cosec}^2 \varphi$ . These deductions accord with experience. Furthermore, the amplitude of the pressure rise would also show a positive correlation with the rainfall, which fact was established by Levine(13). The magnitude of the primary inflow, on the basis of this theory, would be many times greater than the effect computed by Sawyer(12). This, however, does not constitute an objection, for it will be shown in Section V that only a fraction of any such inflow can ultimately be realized as a surface pressure change.

## V. HORIZONTAL ACCELERATION IN THE TRAJECTORY

In the previous section, the importance to the surface pressure change of the last term in equation (7) has been discussed. An earlier paper(1) has dealt with the control of pressure inherent in the first term of the equation. It remains to examine the effect of the middle term  $\frac{\dot{G}}{\lambda} \mathbf{n}_i$ , which represents the gradient component due to horizontal acceleration in the trajectory.

The general effect of this component is to operate against any very rapid change in the surface pressure pattern. Any net flow from high towards low or low towards high pressure will tend to cause respectively a decrease or increase in the pressure gradient: the component  $\frac{\dot{G}}{\lambda} \mathbf{n}_i$  then comes into operation in the opposite sense, and remains until the steady state is once again approached. This exerts a form of restraint supplementary to the "curvature control" which has been the subject of earlier study(1). The two together are the main factors in the general mutual adjustment between air movement and pressure change, the results of which are observed in the high degree of conservatism of both moving and stationary pressure patterns.

As before, it is useful to draw a formal distinction between translation of a pressure system and the changes, development and distortion, which it may experience as it moves. Using  $\frac{\delta}{\delta t}$  to denote the latter we may write, with the same degree of approximation as in equation (7),

$$\dot{G} = -C \frac{\partial G}{\partial x} + \frac{\delta G}{\delta t} + G \frac{\partial G}{\partial s} \dots\dots\dots (12)$$

where  $C$  is the rate of translation of the system eastwards. The first two terms on the right are secondary rather than primary terms. They cannot be determined from the streamline or isobar pattern alone, but implicitly contain some function of the pressure tendency itself. As a result they exert their influence in the form of a restraint on the freedom of translation and distortion respectively. Formally, the terms so arising are absorbed in the control factor  $F$  on the left of equation (1).

An example of the control of distortion, which might suitably have been given here, was included in Section 11 of the earlier paper(1). It now becomes

clear that this case, that of the broad meridional flow pattern, is subject to "isallobaric" rather than "curvature control".

To examine more specifically the direct contribution from the horizontal acceleration, we substitute in and expand the last term of equation (1) by

$$\begin{aligned}
 - \int_0^{p_0} \frac{1}{\rho} \operatorname{div} (\rho \mathbf{A}) dp &= - \int_0^{p_0} \frac{1}{\rho} \operatorname{div} \left( \frac{\rho \dot{G}}{\lambda} \mathbf{n}_i \right) dp \\
 &= - \int_0^{p_0} \left[ \frac{\dot{G}}{\lambda} \operatorname{div} \mathbf{n}_i + \frac{1}{\rho} \frac{\partial}{\partial n_i} \left( \frac{\rho \dot{G}}{\lambda} \right) \right] dp \\
 &= - \int_0^{p_0} \left[ \frac{\dot{G}}{\lambda} \kappa_s + \frac{1}{\rho} \frac{\partial}{\partial n_i} \left( \frac{\rho \dot{G}}{\lambda} \right) \right] dp \quad \dots (13)
 \end{aligned}$$

The divergence given by (13) consists of two terms. The first of these represents the divergence associated with a component everywhere at right angles to a curved isobar, such as P Q R in Figure 3: such directional divergence is termed "diffluence" (14). The second represents a dilatation of air in that the velocity of flow across P Q R may differ from that across a segment of a neighbouring isobar P' Q' R'.

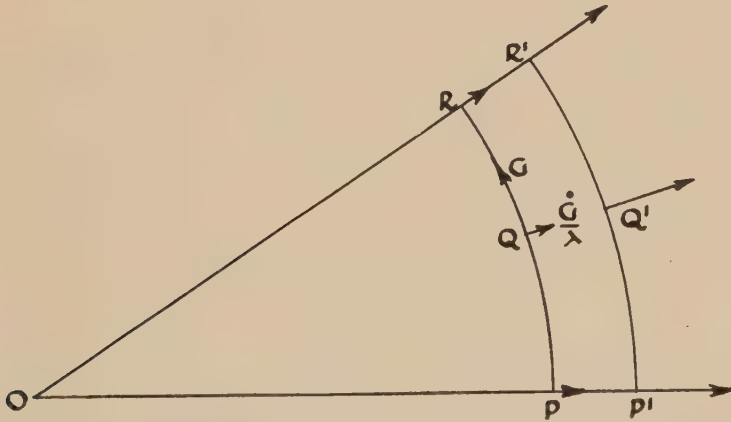


Fig. 3.—Diffluence and dilatation of the deviation due to horizontal acceleration.

Apart from its control effect, the importance of the primary divergence of the horizontal acceleration may be seen from (12) and (13) to lie in the two terms which are vertical integrals of  $\frac{1}{\lambda} \kappa_s G \frac{\partial G}{\partial s}$  and  $\frac{1}{\rho} \frac{\partial}{\partial n_i} \left( \frac{\rho}{\lambda} G \frac{\partial G}{\partial s} \right)$  respectively. Again the overall (development) importance of the latter term may be assessed by smoothing out its irregularities, since the latter can result at most in a local distortion. Since  $G \frac{\partial G}{\partial s}$  is zero at the centre of a closed pressure system, the result of smoothing  $\frac{1}{\rho} \frac{\partial}{\partial n_i} \left( \frac{\rho}{\lambda} G \frac{\partial G}{\partial s} \right)$  is to yield  $\frac{1}{\lambda} \kappa_s G \frac{\partial G}{\partial s}$ , which effectively doubles the contribution to (13) from the confluence or diffluence term. Furthermore

by reference to equation (1) it is seen that this term is already explicit in the complete tendency equation, and the overall effect of the horizontal acceleration on the primary pressure change is that the contribution from the diffluence  $\frac{1}{\lambda} \kappa_s G \frac{\partial G}{\partial s}$  occurs with the factor 3 instead of unity. The practical significance of this term has already been fully discussed in the earlier paper(1) and, in the typical pressure systems there brought under examination, the movement was importantly affected by this term in only one instance (conclusion 3). The modified conclusion is that, other things being equal, cold "lows" and warm "highs" will tend to move in the direction of the strongest wind around them, at a speed of the order of the difference between the greatest and least mean\* circumferential wind.

With symmetrical systems the term  $\frac{1}{\lambda} \kappa_s G \frac{\partial G}{\partial s}$  will be of little importance to development. Neglecting the variations of  $\kappa_s$  and  $\lambda$ , the integral round a closed circuit is zero, and in such circumstances there can be no overall cyclo- or anti-cyclogenesis. The importance of the term in cases of asymmetry can be

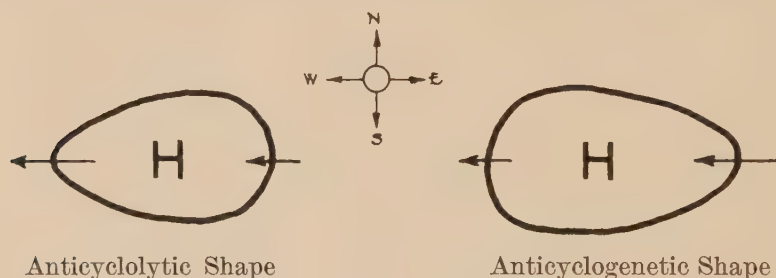


Fig. 4.—Arrows show the horizontal acceleration deviation.

illustrated by the example of the anticyclone. Owing to the fall of temperature with increasing latitude, the greatest mean\* circumferential wind in either hemisphere is commonly found on the poleward side. In either hemisphere, then, the horizontal acceleration deviation is directed from east to west on both sides of the anticyclone. There will therefore be an anticyclogenetic influence when the curvature is greater east of the centre, and an anticyclolytic tendency when the curvature is greater west of the centre. The extent to which this influence can make itself felt is a matter for synoptic investigation. It may well provide the explanation for the eastwards "thrust of a sharp ridge" frequently observed in the North Atlantic Ocean and the western half of the Australian continent.

Finally, the irregular part of  $\frac{1}{\rho} \frac{\partial}{\partial n_i} \left( \frac{\rho G}{\lambda} \right)$  appears to be of potential significance only in distorting the isobar pattern. Consisting as it does of a second derivative

\* By "mean" is implied the average, weighted proportionately to pressure, from the surface to the limits of atmosphere above the point in question.



of the wind, or a third space (or two space—one time) derivative of the pressure, it is impossible with present station networks (even where as closely spaced as in Britain or the United States) to assess its importance in individual cases. Subject to a further advance in knowledge or instrumental techniques it must remain as the imponderable factor in the problem of pressure variations. It is hard to conceive a natural flow pattern in which the irregularities can operate systematically over a sufficiently large area of the system to alter radically the translation of the pressure system as determined by the factors already considered. This is a qualitative or instinctive argument beyond which the claim, or hypothesis, to neglect the irregular part of  $\frac{1}{\rho} \frac{\partial}{\partial n_i} \left( \frac{\rho \dot{G}}{\lambda} \right)$  in broad pressure changes cannot be substantiated at present. If it proves to be unfounded, the very nature of the term implies that the prospect of a complete understanding of surface pressure variations is still distant.

## VI. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Section of Meteorological Physics, C.S.I.R.

## VI. REFERENCES

- (1) PRIESTLEY, C. H. B.—*Quart. J. R. Met. Soc.* **73**: 65 (1947).
  - (2) PRIESTLEY, C. H. B.—*Nature* **158**: 914 (1946).
  - (3) SUTCLIFFE, R. C.—*Quart. J. R. Met. Soc.* **64**: 495 (1938).
  - (4) DURST, C. S., and SUTCLIFFE, R. C.—*Ibid.* **64**: 75 (1938).
  - (5) PALMER, C. E.—N.Z. Met. Office, Prof. Note No. 2 (1942).
  - (6) BRUNT, D.—“Physical and Dynamical Meteorology”. 2nd Ed. (Cambridge Univ. Press, 1939.)
  - (7) DURST, C. S.—*Quart. J. R. Met. Soc.* **59**: 231 (1933).
  - (8) HANN-SÜRING.—“Lehrbuch der Meteorologie.” 5th Ed. Vol. 8, pp. 707-20 (1942).
  - (9) ELSASSER, W. M.—Harvard Met. Stud. No. 6 (1942).
  - (10) THOMAS, H.—*S.B. preuss. Akad. Wiss.* **17** (1934).
  - (11) VAN MIEGHAM, J.—*Inst. R. Met. Belg., Mem. No.* 7 (1937).
  - (12) SAWYER, J. S.—*Quart. J. R. Met. Soc.* **72**: 168 (1946).
  - (13) LEVINE, J.—*Bull. Amer. Met. Soc.* **23**: 58 (1942).
  - (14) HOGGEN, G. L.—*Quart. J. R. Met. Soc.* **72**: 319 (1946).
-

# OBSERVATIONS ON THE VARIABLE SOURCE OF COSMIC RADIO FREQUENCY RADIATION IN THE CONSTELLATION OF CYGNUS

By J. G. BOLTON\* and G. J. STANLEY\*

(Plates 1 and 1A)

[Accepted for Publication January 16, 1948]

## *Summary*

An account is given of investigations on a strong variable source of radio frequency energy in the constellation of Cygnus, including a determination of the size and celestial coordinates of the source and a study of the variations over a frequency range from 60 to 200 Mc/s. The frequency spectra are compared with those of galactic and solar radiation.

The position of the source is found to be in a region of relatively low stellar density and not coinciding with an outstanding stellar object. A discussion of limits of distance and possible nature of the source is given.

## I. HISTORICAL SURVEY

Since the discovery of cosmic or galactic noise by Jansky in 1931(1) various workers have made surveys of the intensity of this noise over the celestial sphere, covering a wide range of frequencies. The origin of galactic noise is as yet unknown; it may be due to free-free electron transitions in ionized interstellar matter, or enhanced radiation from individual stars similar to that received from the sun during times of sunspot activity—or a combination of both.

Probably the most accurate plot yet made of galactic noise is due to Hey, Parsons, and Phillips(2). They observed at a frequency of 64 Mc/s., and, using a fairly narrow beam aerial and a novel method of analysis, found the greatest intensity of noise in the Scorpio-Sagittarius region and a second sharply defined peak in the constellation of Cygnus. Later they(3) showed that the noise level from Cygnus was by no means constant, and that variations with a period of about one minute occurred. They found that the variations were neither of solar nor terrestrial origin, and estimated that the angular width of the variable source was less than  $2^\circ$  and centred at Right Ascension 20 hours, Declination  $43^\circ$ .

During the past three months a special study has been made of this region at frequencies of 60, 85, 100, and 200 Mc/s. The technique employed was that first developed by Pawsey, Payne-Scott, and McCready(4) for studying solar radiation at 200 Mc/s., i.e. the receiving equipment is situated on a high cliff and the source observed as it rises over the sea. The direct ray and the ray reflected

\* Division of Radiophysics, C.S.I.R.

from the sea alternately cancel and reinforce as the path difference varies between even and odd multiples of  $\lambda/2$ , giving rise to a succession of maxima and minima.

From the times of occurrence of these maxima and minima the track of the source above the horizon can be deduced, and hence its celestial coordinates determined. An estimate of the size of the source may be made from the ratio of intensities of the maxima and minima.

## II. EQUIPMENT AND SENSITIVITIES

Table 1 gives a summary of the characteristics of the aerial systems and receivers used in the present investigation. The figures in the last column represent the intensity of a received signal equivalent to the receiver fluctuation level in watts  $m^{-2}$  (c/s.)<sup>-1</sup>.

TABLE 1  
AERIAL AND RECEIVER DETAILS

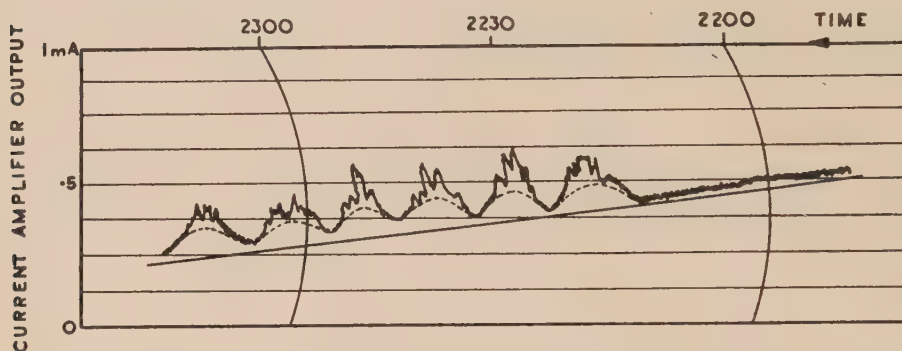
Fre- quency  (Mc/s.)	Aerial Array	Effective Area  (sq. metres)	Receiver		Fluctuation Level as a Fraction of Total Re- ceiver Noise	Input Signal equivalent to Receiver Fluctuation Level (Watts $m^{-2}(c/s.)^{-1}$ )
			Noise Factor	Band- width  (Mc/s.)		
60	2 Yagis, $\lambda/2$ apart, in ver- tical plane	12	3	1.5	1/1000	$2 \times 10^{-24}$
85	1 Yagi	5	5	1.5	1/2000	$4 \times 10^{-24}$
100	2 Yagis, $\lambda/2$ apart, in ver- tical plane	8.5	6	1.5	1/2000	$3 \times 10^{-24}$
200	4 Yagis, $\lambda$ apart, in square array	4.3	8	1.5	1/1000	$1.8 \times 10^{-23}$

The aerials consisted of one or more Yagis and were carefully matched to the receivers, the standing wave ratios being less than 1.1/1. The receivers were converted radar receivers, with no notable modifications other than the provision of a two-stage preamplifier in the 200 Mc/s. set. For recording small signals a stage of current amplification was added after the detector in all cases. This, in effect, balanced out the noise generated in the receiver and further amplified the input signal for presentation on a recording milliammeter.

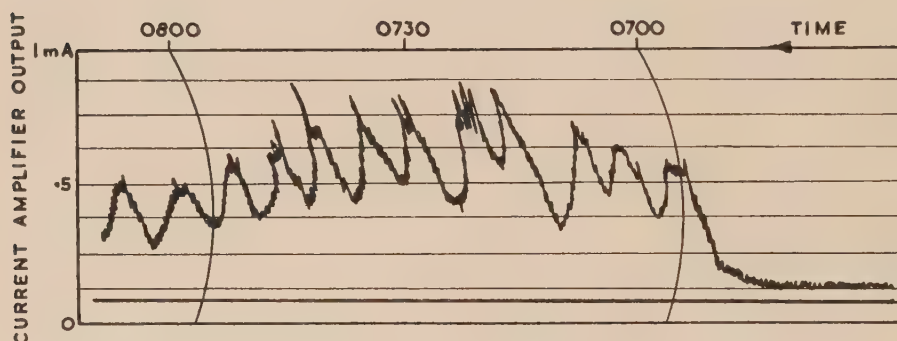
Calibration of the records was achieved by switching-in a matched load in place of the aerial at desired intervals.

## III. GENERAL RESULTS

Figure 1 (A) shows a typical record obtained on a frequency of 100 Mc/s. when the region under consideration rises over the sea. The gradual decline in



A



B

Figs. 1 (A) and 1 (B).—Typical noise level records showing interference between the direct and reflected rays as a source rises over the sea. Frequency 100 Mc/s. Height of aerial above mean sea-level 288 ft.

A. Variable source in Cygnus, June 19, 1947.

B. The sun, June 24, 1947.

noise level throughout the record is due to the intense region of cosmic noise in Sagittarius passing out of the aerial beam. Three facts are apparent :

- (i) The existence of two components, one constant and one variable. A dotted sine curve has been drawn on the figure roughly separating the two components.
- (ii) The great depth of the minima and absence of fluctuations in the minima, indicating a source small compared with the width of the interference lobes. A pattern from the “quiet” sun rising is given for comparison in Figure 1 (B), showing the shallow minima obtained from a spread source.



- (iii) The intensity of the noise from this region represents a considerable fraction of the total noise received by the wide-beam aerials, and is comparable to that of the "quiet" sun.

Similar records are obtained on 60, 85, and 200 Mc/s. except that on the two lower frequencies the variable component tends to overshadow the constant component, and on 200 Mc/s. only a constant component has been observed.

The majority of our results have been obtained at a frequency of 100 Mc/s. This frequency was preferred because of freedom from man-made interference and because the aerial array was compact and suitable for field observations.

#### IV. THE TWO COMPONENTS

On referring again to Figure 1 (A) it will be seen that a sine curve has been drawn through the lower limits of the fluctuations. Comparison of a large number of records reveals that the intensity enclosed by this curve remains constant within the limits of experimental error and reading accuracy. Table 2

TABLE 2  
INTENSITY OF SOURCE

PART I.—LEVELS OF CONSTANT AND VARIABLE COMPONENT

Date	Constant Component (100 Mc/s.) (watts $m^{-2}$ (c/s.) $^{-1}$ )	Mean Intensity of Variable Component
June 19 .. ..	$6 \times 10^{-23}$	$3 \times 10^{-23}$
June 20 .. ..	$6 \times 10^{-23}$	$1 \times 10^{-23}$
June 24 .. ..	$*5.5 \times 10^{-23}$	$3 \times 10^{-23}$
July 3 .. ..	$*5 \times 10^{-23}$	$3 \times 10^{-23}$
July 15 .. ..	$6 \times 10^{-23}$	$6 \times 10^{-23}$
July 17 .. ..	$*5.5 \times 10^{-23}$	nil
July 25 .. ..	$6 \times 10^{-23}$	$4 \times 10^{-23}$
July 26 .. ..	$6 \times 10^{-23}$	$3 \times 10^{-23}$
July 28 .. ..	$6 \times 10^{-23}$	$3 \times 10^{-23}$
August 15 .. ..	$6 \times 10^{-23}$	nil
September 6.. ..	$6 \times 10^{-23}$	$1 \times 10^{-23}$

\* Deviation within limits of reading accuracy of the relevant records.

PART II.—VARIATION OF CONSTANT COMPONENT WITH FREQUENCY

Frequency	200 Mc/s.	100 Mc/s.	85 Mc/s.	60 Mc/s.
Intensity in watts $m^{-2}$ (c/s.) $^{-1}$ ..	$3 \times 10^{-23}$	$6 \times 10^{-23}$	$5 \times 10^{-23}$	$4 \times 10^{-23}$

is a sampling of results obtained between June and September 1947. On a number of occasions the variable component was absent altogether, making measurement of the intensity of the constant component more certain. So far no variable component has been observed on 200 Mc/s., and it has been found

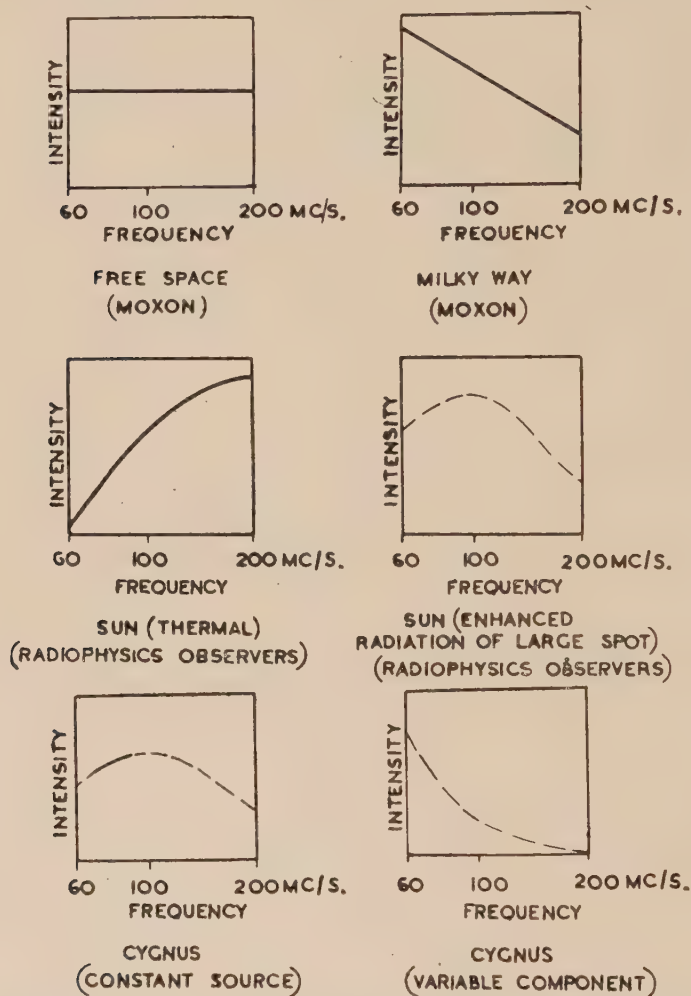


Fig. 2.—Qualitative comparison of spectra of solar and galactic noise.

that the intensity of a variable component and the possibility of its presence increase with decreasing frequency. The spectrum of the constant component appears to have a shallow peak in the region of 100 Mc/s. whereas the amplitude of the variable component increases sharply at lower frequencies. Qualitative spectra of the two components are shown in Figure 2 together with the spectra of thermal and enhanced radiation from the sun and of galactic noise.

## V. SIZE OF SOURCE

An upper limit may be placed on the size of the source by comparing the relative heights of the maxima and minima resulting from interference between the direct ray and the ray reflected from the sea. In the ideal case of a receiver with infinitely small bandwidth, and regarding the sea as a perfect plane reflector, it can be shown that the received power from a point source varies as

$$\sin^2 \frac{2\pi h\alpha}{\lambda}$$

where  $h$  = height of aerial above sea

$\alpha$  = elevation of source

$\lambda$  = wavelength.

Maximum and minimum values occur for  $2\pi h\alpha/\lambda = (2n+1)\pi/2$  and  $n\pi$ .

For a "spread" source the minima are not zero and the maxima are smaller than in the ideal case. If the radiation originates in a uniformly radiating strip

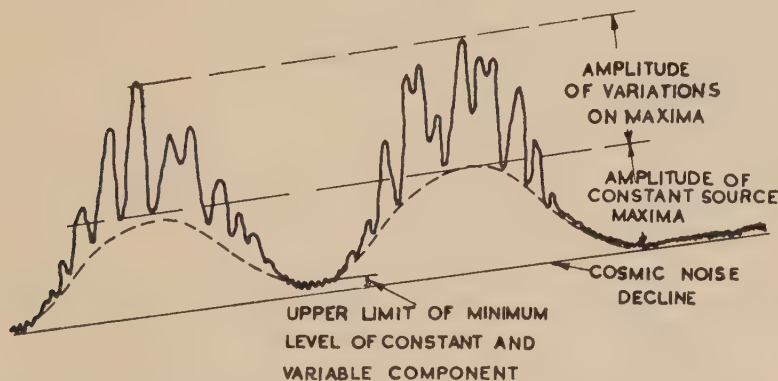


Fig. 3.—Enlarged section of a record of noise from Cygnus showing how maxima and minima of the constant and variable components are identified.

of angular width  $W$  rising parallel to the horizon, it can be shown that  $W$  is related to the ratio  $R$  of the height of the minima to that of the maxima by the formula.

$$W = \frac{\lambda}{\pi h} \sqrt{3R}.$$

In the case of the variable component,  $R$  is the ratio of the relative amplitudes of the variations on the maxima to those on the minima. The latter were found to be very small—less than the receiver fluctuation level—so that inherent receiver stability placed an upper limit on  $R$ . Under the best conditions, values of  $R$  greater than 50 were found, giving a width of the radiating strip of less than  $8'$ .

For the constant component the base level from which the heights of maxima and minima were reckoned was the extrapolated decline of cosmic noise (see Fig. 3). If we assume this extrapolation is reliable for the first few lobes, the same limit as above is found to apply in the case of the height of the minima.

Thus only an upper limit could be placed on the width of the radiating strip—of the same order as for the variable component.

In the course of the source's path above the horizon, the orientation of the original radiating strip with respect to the horizon changed by more than  $90^\circ$ . From the absence of variations in the minima at all positions, the upper limit to the size of the source can be said to have circular symmetry.

Although variations in the minima can only be said to be less than the receiver fluctuation level, the receiver fluctuations have a much shorter period than the variations of the source. Thus, careful examination of the records suggests a much smaller source size than stated above. Further experiments using improved receiver stability and greater aerial height will probably substantiate the authors' belief that the source is effectively a "point".

## VI. CORRELATION OF VARIATIONS AT DIFFERENT FREQUENCIES

Study of the behaviour of the variable component at different frequencies was somewhat difficult owing to the use of the interference technique, and only limited results over short periods were obtained.

Results so far indicate that:

- (i) There is little or no correlation between individual variations on different frequencies, but there is fair correlation between the mean amplitude of variations on different nights. This is similar to the behaviour of enhanced solar radiation.
- (ii) Variations are more likely to occur on the lower frequencies, i.e. while there may be no variations on 100 Mc/s., 85 Mc/s. may show slight activity and 60 Mc/s. considerable activity. This differs from solar enhanced radiation over this range of frequencies where, particularly in the case of small sunspots, disturbances occur on 100 Mc/s. but not on the lower frequencies.
- (iii) If simultaneous observations are made on two frequencies, the variations on the lower frequency have a shorter period.
- (iv) On any one frequency the variations of greater amplitude seem to have shorter periods. The overall range of periods is 0.1 to 1 minute. Typical records of fast and slow variations are shown in Figures 4 (A) and 4 (B).

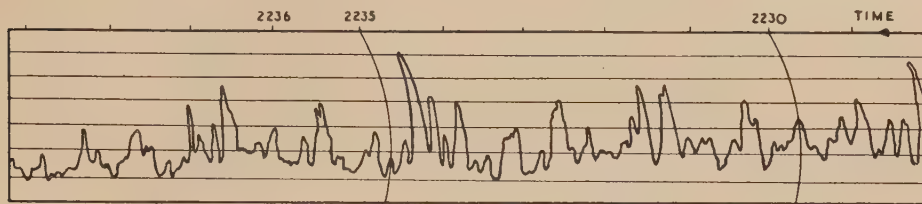
These results are based on a limited number of records and require statistical confirmation.

## VII. LOCATION OF THE SOURCE

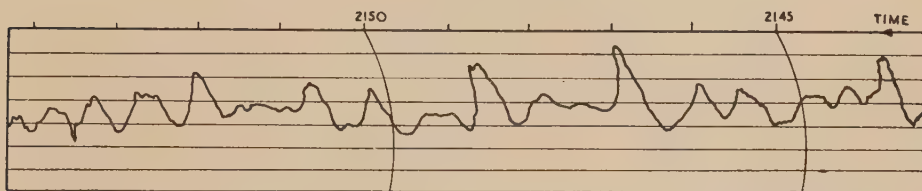
Location of the source was achieved by using the times of occurrence of the minima of a lobe pattern during the complete passage of the source above the horizon. Right Ascension, i.e. the sidereal time of culmination, was derived from times of minima at equal elevations on either side of the meridian. From the time of culmination and the estimated time of rising, the semi-diurnal arc, and thus declination, was obtained.



A single site was not available in New South Wales possessing the two necessary features—a clear view over the sea from rising to setting of the source, and sufficient height to give a large number of minima for accurate timing. Location was therefore carried out in four stages.



A



B

Figs. 4 (A) and 4 (B).—Typical noise records, taken on a frequency of 100 Mc/s., height of aerial above mean sea-level 100 ft.

A. Rapid variations, July 27, 1947.

B. Slow variations, July 25, 1947.

- (i) *Approximate Right Ascension*.—Observations were made from a low headland with the necessary clear view. Only five minima were obtained from rising to culmination and five from culmination to setting, and these suffered displacements in time due to tidal changes in sea-level over that period. However, a mean of observations over five nights gave a Right Ascension correct to  $\pm 1$  minute.
- (ii) *Accurate Right Ascension*.—Observations were made from a high headland overlooking a river estuary. Opposing cliffs and an island obscured the lobe pattern except for a short time on either side of culmination. The approximate Right Ascension previously determined enabled identification of the corresponding minima to be made. A sample series of these, obtained when high tide coincided with culmination, is shown in Table 3. The maximum possible shift of one set of minima with respect to the other due to tidal changes in sea-level was 5 seconds, giving an error in time of culmination of 2.5 seconds. Reading accuracy of any one minimum was 5 seconds.

TABLE 3

	Time of Corresponding Minima (Eastern Australian Standard Time)	Time of Culmination
A A'	22 hr. 10·8 min. } 03 22·4 }	00 hr. 46·6 min.
B B'	22 03·3 } 03 30·3 }	00 46·8
C C'	21 55·0 } 03 38·5 }	00 46·75
D D'	21 48·0 } 03 45·5 }	00 46·75
	Mean Time of Culmination	00 hr. 46·73 min.

Results obtained on July 10, 1947. Frequency 100 Mc/s. Height above mean sea-level 390 ft. Longitude of site  $151^{\circ} 18' 40''$ .

(iii) *Declination*.—The path of the rising source was reconstructed to find the actual time of rising. Each minimum, representing an apparent elevation, was corrected for refraction and earth's curvature to give the true elevation at that time. A curve of elevation against time was drawn for a number of nights and the mean sidereal time of rising found. The declination was then calculated from the latitude of the site and the semi-diurnal arc.

(iv) *Check of Declination*.—In order to improve the accuracy and avoid relying on extrapolation of a curve, a further method was used to check the declination. The relation between the elevation of the source  $\alpha$ , the latitude of the site  $l$ , the declination of the source  $\delta$ , and the hour angle  $\theta$ , is

$$\sin \alpha = \cos l \cos \delta \cos \theta - \sin l \sin \delta.$$

Assuming a value of  $\delta$  where  $\delta + \Delta$  is the true value, if  $\Delta$  is a small angle it can be shown that

$$\Delta = \frac{\sin \alpha + \sin l \sin \delta - \cos \theta \cos l \cos \delta}{\cos \theta \cos l \sin \delta + \sin l \cos \delta}.$$

Each minimum of a particular record gave a value of  $\Delta$  and a statistical mean was obtained, neglecting readings obviously influenced by poor reading accuracy. The position of the source obtained by these methods is:

$$\begin{array}{lll} \text{Right Ascension} & 19 \text{ hr. } 58 \text{ min. } 47 \text{ sec.} & \pm 10 \text{ sec.} \\ \text{Declination} & +41^{\circ} \quad 47' & \pm 7'. \end{array}$$

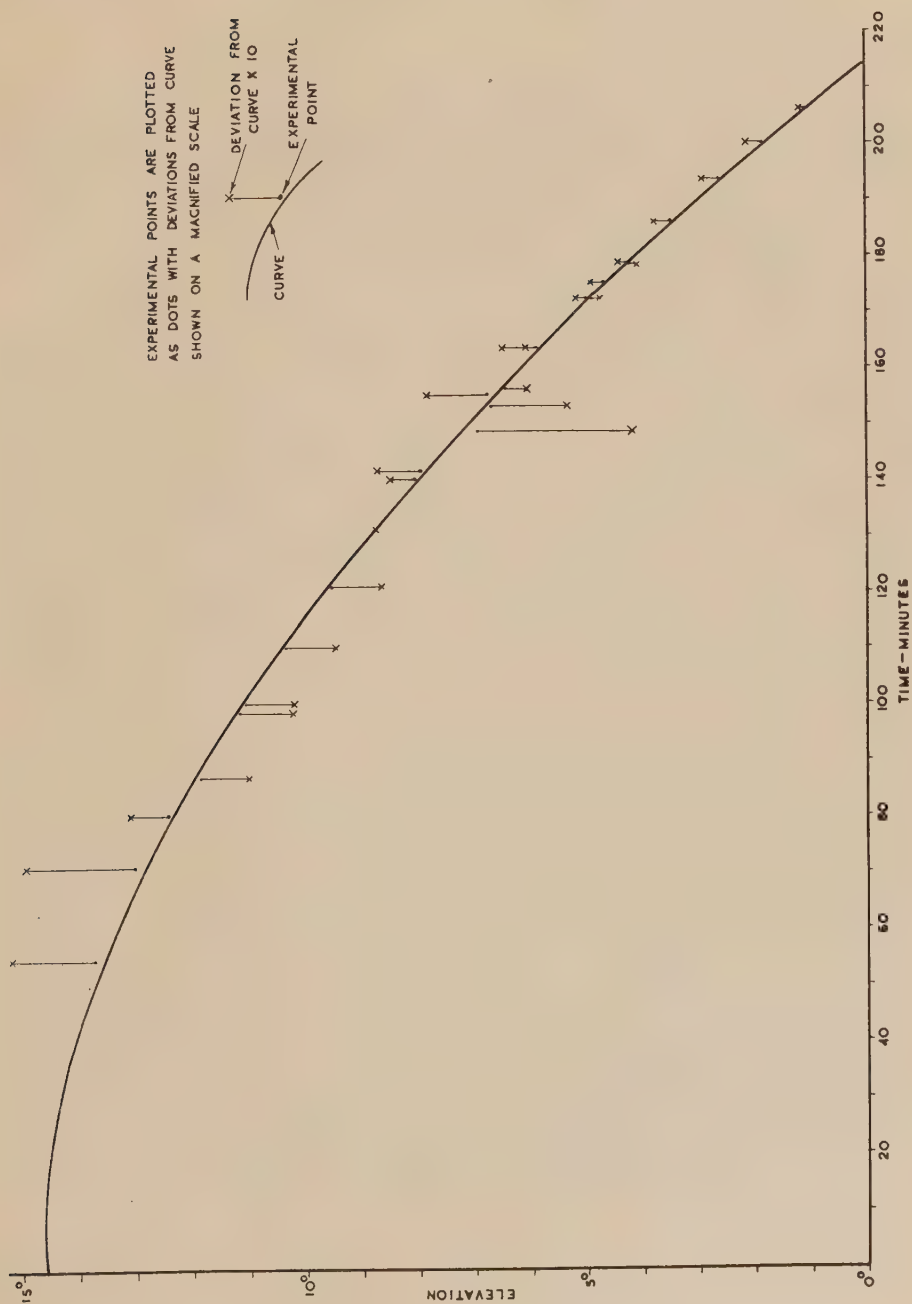


Fig. 5.—Path of a source at declination  $+41^{\circ} 47'$  as seen from latitude  $33^{\circ} 34' S$ .

Figure 5 shows the path of the source for latitude  $33^{\circ} 34' \text{ S.}$  and declination  $+41^{\circ} 47'$ . Experimental points for two sets of readings are plotted to show the scatter. The dots indicate the determined points and the crosses show the deviation of these points from the curve magnified ten times.

### VIII. DISCUSSION OF RESULTS

A powerful source of radio frequency energy has been studied and the following data established :

(1) Its position is :

Right Ascension 19 hr. 58 min. 47 sec.  $\pm 10$  sec.

Declination  $+41^{\circ} 47'$   $\pm 7'$ .

(2) Its angular width is less than  $8'$  of arc.

(3) It has two components, one believed constant, and the other showing considerable variations with time.

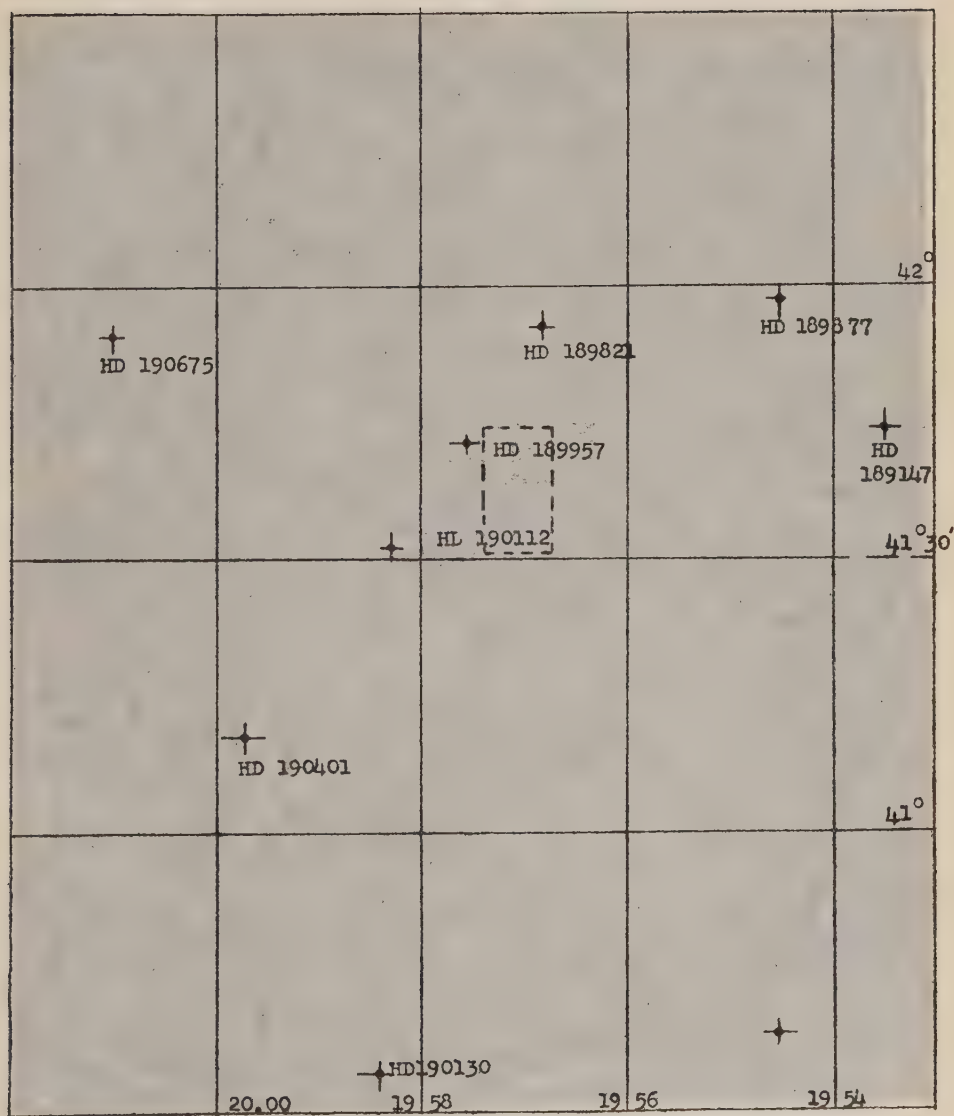
Reference to star catalogues, in particular the Henry Draper Catalogue, shows that the source is in a region of the galaxy distinguished by the absence of bright stars and objects such as nebulae, double and variable stars, i.e. the radio noise received from this region is out of all proportion to the optical radiation. Although the experimental technique allows only an upper limit to be placed on the size of the source, this is believed to be effectively a point and therefore a single object. The determined position lies in a less crowded area of the Milky Way and the only obvious stellar objects close to the stated limits of accuracy are two seventh magnitude stars. There is certainly no comparable optical radiation from this region.

The radiation from this source probably accounts entirely for the secondary maximum in galactic noise noted by Hey, Phillips, and Parsons. Observation of this region of the galaxy rising above the sea shows only a discrete source superimposed on a smooth continuum. At rising, the constant component on 100 Mc/s. is about one-twentieth of the total noise received by our aerial system. For an area of  $8'$  angular width, the noise brightness is about 1500 times that of the general background. Hey, Phillips, and Parsons' contour map of galactic noise at 64 Mc/s. shows an area of some  $3^{\circ}$  angular width with twice the noise brightness of the background ; for an area of  $8'$  the equivalent brightness would be about 500, showing qualitative agreement between the two results.

So far there is little evidence from which an estimate of the distance of the source may be made. Although the location differs from that of Hey by  $1^{\circ}$  in declination, the variation of the estimated time of rising over the three months is smaller than the reading accuracy of 10 seconds, i.e. the corresponding parallax is less than  $2\frac{1}{2}''$  of arc. Also the discrepancy of  $1^{\circ}$  is within the limits of Hey's determination of position. Two hypothetical limits on the distance may be suggested.

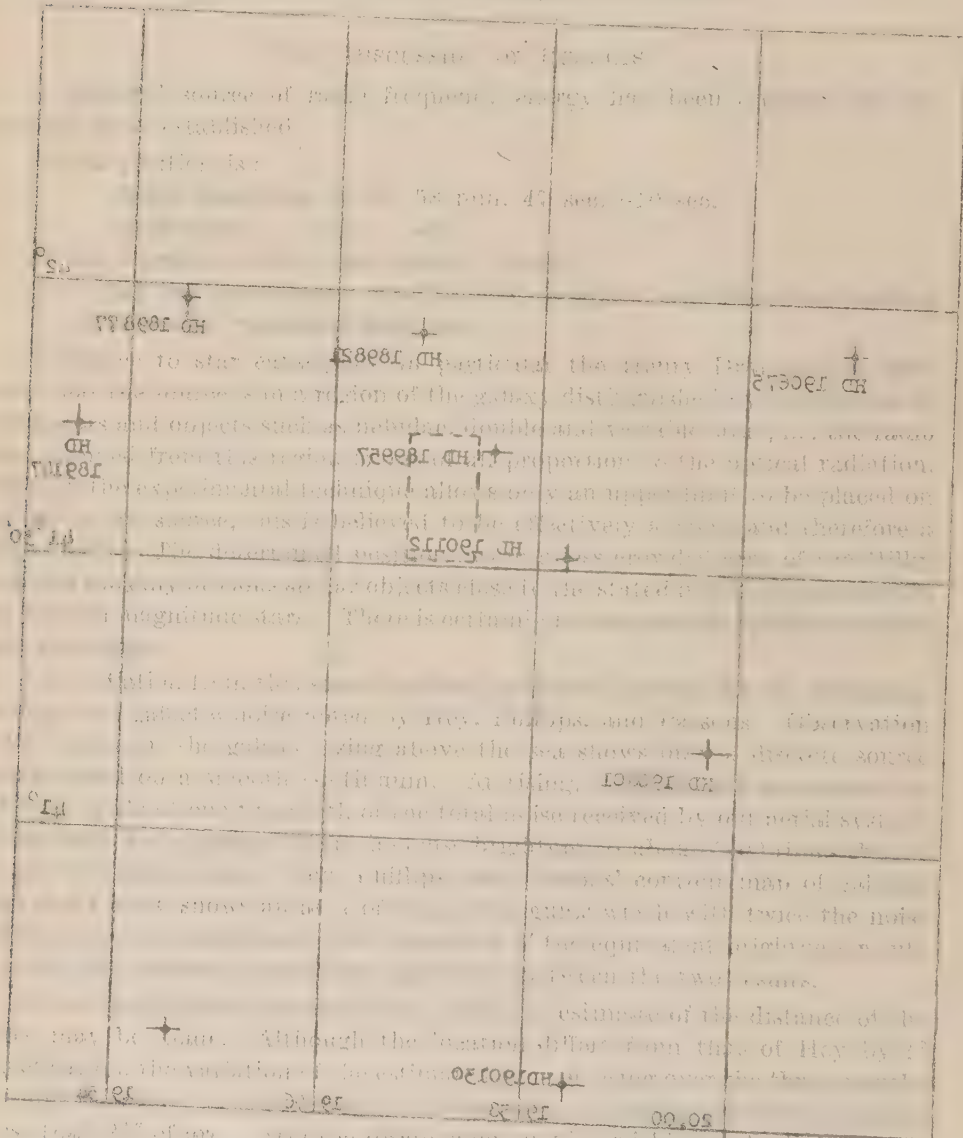
- (i) *A Lower Limit of Distance.*—If the variations extend over the whole source it is unlikely that the width of the source would exceed the product of the velocity of light and the period of the shortest variation. Considering the angular width as  $8'$  and a period of variation of 0.25





BOLTON and STANLEY.—OBSERVATIONS ON THE VARIABLE SOURCE OF COSMIC RADIO  
FREQUENCY RADIATION IN THE CONSTELLATION OF CYGNUS

and the 1000 Mc. band is the only one in which the signal is not masked by the noise of the atmosphere. The signal is a continuous wave of constant frequency and amplitude, and is not affected by the ionosphere. It is the only one of its kind in the world.



THE VARIATION OF COSMIC RAY RADIATION IN THE CONSTITUTION OF COSMOS  
HOLTON AND STANLEY.—OBSERVATIONS ON THE VARIABLE SOURCE OF COSMIC RADIATION  
The cosmic ray radiation is a continuous wave of constant frequency and amplitude, and is not affected by the ionosphere. It is the only one of its kind in the world.



BOLTON and STANLEY.—OBSERVATIONS ON THE VARIABLE SOURCE OF COSMIC RADIO  
FREQUENCY RADIATION IN THE CONSTELLATION OF CYGNUS





minute the distance of the source cannot exceed 100 light-minutes. This estimate would give an easily discernible parallax. However, it has been pointed out above that the size of the source is probably considerably less than  $8'$  and there is no reason to suppose that the variations extend over the whole surface.

- (ii) *An Upper Limit of Distance.*—Assume the source to be an average star but with its total energy output in the radio frequency spectrum. Taking a mean intensity of  $2 \times 10^{-23}$  watts  $m^{-2}$   $(c/s.)^{-1}$  for the source spread over a bandwidth of 1000 Mc/s. the total radiation would be  $2 \times 10^{-14}$  watts  $m^{-2}$ . The sun considered as an average star has a total radiation of about 1 kilowatt per square metre. Applying an inverse square law this gives the distance of the source as some 3000 light-years.

If we consider the size of the source to be less than  $8'$ , the effective temperature at 100 Mc/s. would be greater than  $4 \times 10^6$  deg. K. This makes a thermal origin of the radiation improbable. However, a mechanism similar to that proposed by Martyn(5) to account for the steady enhanced noise from a large active sunspot—perhaps the association of moving ionized matter and strong magnetic fields—is quite possible. A possible cyclic occurrence and the state of polarization of the variable component may yield valuable additional information. Observations in progress at the time of writing reveal that the variable component may be absent for as long as two weeks over the range of frequencies considered. In view of the probability of further information being obtained on the behaviour and characteristics of this source, the question of its nature is left open for the present.

#### IX. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.

#### X. REFERENCES

- (1) JANSKY, K. G.—Directional studies of atmospherics at high frequencies. *Proc. Inst. Radio Engrs.*, N.Y. **20**: 1920-32 (1932).
- (2) HEY, J. S., PHILLIPS, J. W., and PARSONS, S. J.—Cosmic radiation at 5 metre wavelength. *Nature* **157**: 296-7 (1946).
- (3) HEY, J. S., PHILLIPS, J. W., and PARSONS, S. J.—Fluctuations in cosmic radiation at radio frequencies. *Ibid.* **157**: 805-6 (1946).
- (4) MCCREADY, L. L., PAWSEY, J. L., and PAYNE-SCOTT, R.—Solar radiation at radio frequencies and its relation to sunspots. *Proc. Roy. Soc. A* **187**: 357 (1947).
- (5) MARTYN, D. F.—Origin of radio emissions from the disturbed sun. *Nature* **159**: 26-7 (1947).

#### EXPLANATION OF PLATES 1-1A

Star-field in the constellation of Cygnus. The position of the variable source of radio frequency energy is shown, within the upper limits of experimental error, by the dotted rectangle. Of the two "bright" stars close to this area HD189957 is of magnitude 7.6, spectral type B-3 and HD190112 7.2, type A-O (Epoch 1900).

[Photograph by courtesy of the Commonwealth Observatory, Mount Stromlo, Canberra.]

# THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY CONTAINING TWO PHASES

By R. W. K. HONEYCOMBE\* and W. BOAS\*

(Plates 1-8)

[Accepted for Publication October 28, 1947]

## *Summary*

The deformation and recrystallization of a duplex brass containing 60 per cent. copper and 40 per cent. zinc has been studied using microscopic and X-ray methods. In compression and tensile tests the  $\alpha$  phase commences to deform plastically at a lower stress than the  $\beta$  phase, and even after heavy deformations, it is deformed more severely. Which of the two phases starts to recrystallize first on annealing depends on the rate of cooling of the alloy prior to deformation.

In the quenched wires the  $\alpha$  phase commences to recrystallize about 100° C. below the  $\beta$  phase, whereas in the slowly cooled wires the  $\beta$  phase commences to recrystallize before the  $\alpha$  phase. In the former case, the  $\beta$  recrystallizes at a lower temperature, the smaller the amount of  $\beta$  present, i.e. the lower the temperature of quenching. It has been shown that the order-disorder transformation is responsible for the difference of behaviour of the slowly cooled wires from that of the quenched wires.

Microscopic examination shows that the  $\alpha$  phase starts to recrystallize at random whereas the  $\beta$  phase tends to nucleate at the boundaries between the two phases. The experiments indicate that it is not possible to conclude that a phase has suffered more severe deformation because it commences to recrystallize at a lower temperature.

## I. INTRODUCTION

The deformation and recrystallization of pure metals and alloys containing only one phase are familiar processes which have been very thoroughly studied. On the other hand, little attention has been given to the deformation of alloys containing more than one phase, and we have no knowledge of any work on the recrystallization of such alloys.

Unkel(1) has studied the deformation of a number of two-phase alloys by rolling. He has pointed out that in such alloys the phases will possess different physical properties, for example, yield points and work hardening characteristics, which define the behaviour of the phase during deformation. Unkel has classified the duplex alloys according to whether the second phase is harder or softer than the matrix, and whether it is brittle or ductile. He followed the relative deformation of the two phases by measuring macroscopically the mean compression of a very large number of grains, the section of observation being at right angles to the rolling plane. In alloys in which the second phase is softer and more ductile than the matrix, the deformation of the latter is somewhat less than that of the inclusions. In the initial stages of the deformation, such behaviour would be expected as the softer phase (viz. the phase with the lower yield point) will plastically deform at a lower stress. As the stress is further increased, the yield point of the harder phase will be reached and it will now also

\* Section of Tribophysics, C.S.I.R.

deform plastically. The relative work hardening characteristics of the two phases will decide whether the two phases will continue to be deformed to different extents or whether the difference will be reduced. The latter will occur if the originally softer phase has a much steeper work hardening curve than the harder phase. In the extreme case, the softer phase work hardens so rapidly that it would eventually deform less than the initially harder phase.

This behaviour is further complicated by two factors. Firstly, if in the early stages of the deformation, only one phase deforms, then high stresses will be set up at the junctions of the two phases. Thus additional deformation will occur in the softer phase near these junctions in order to relieve the stresses. This additional flow has been shown to occur by Unckel(1) and by Rhines and Ward(2). Secondly, the yield point and work hardening characteristics of a crystal embedded in crystals of another phase may be quite different from those of the crystal surrounded by crystals of its own type.

There are various grounds on which one would expect the two phases of a deformed duplex alloy not to recrystallize simultaneously. First, the two phases have been deformed to different extents and thus, *ceteris paribus*, they will commence to recrystallize at different temperatures and after different total deformations of the specimen. Further, since the phases are of different chemical composition, they will inherently possess different recrystallization characteristics. Finally, the crystal structure of the phases may play some part, although the rôle of crystal structure on diffusion and nucleation is still unknown.

This paper describes an investigation on the deformation and recrystallization of  $\alpha$ - $\beta$  brass containing approximately 40 per cent. of zinc. The two phases in this alloy are both ductile. The face-centred cubic  $\alpha$  phase is much softer than the body-centred cubic  $\beta$  phase, although the chemical compositions of the two phases are very similar. As this type of alloy has found wide technical application, details of the equilibrium diagram and the properties of the phases are well known. It was thought that the behaviour of this alloy would be typical of a large class of duplex alloys to which the results of this investigation would be applicable.

## II. EXPERIMENTAL

In the majority of the experiments either extruded brass strip 1 in. by  $\frac{1}{4}$  in. or brass rod  $\frac{3}{16}$  in. in diameter was employed. The analyses of these materials were as follows:

	Brass Strip	Brass Rod
	%	%
Copper .. .. .	60.4	60.4
Lead .. .. .	0.02	0.03
Tin .. .. .	Trace	Trace
Nickel .. .. .	Trace	Trace
Iron .. .. .	0.02	0.02
Manganese .. .. .	Trace	Trace
Zinc (balance) .. .. .	39.56	39.55

The treatment of these materials prior to deformation varied considerably and will be described later.

Tensile and compression specimens were machined from the brass strip. The tensile specimens were of the conventional shape, while the compression specimens were  $\frac{1}{2}$  in. by  $\frac{1}{2}$  in. by  $\frac{1}{4}$  in. In the majority of the experiments on recrystallization the brass rod was employed. After annealing it was cold drawn through steel dies to the required reduction in area, employing in all cases the same sequence of dies. Stearic acid was used as a lubricant.

For studying the recrystallization, specimens of wire  $\frac{1}{2}$  in. to 1 in. long were employed. The specimens were annealed by direct immersion in a molten salt bath for 30 minutes. Two identical specimens were annealed together at each temperature, one being used for metallographic examination, the other being etched to 0.025-0.030 in. diameter for X-ray examination.

To facilitate microscopic examination, the small wire specimens were individually mounted in leucite (the temperature of mounting did not exceed 150° C.). A small hole was then drilled through the back of the mounting and electrical contact made on the back of the specimen by means of a piece of copper wire. After preliminary mechanical grinding, the wires were electrolytically polished in orthophosphoric acid (1000 g./litre), the potential difference across the cell being 1.2-1.6 volts. The two phases could be distinguished in the unetched specimens as the natural colours of the phases were not masked by a deformed surface layer. Chemical etching was used to develop the detailed structure of the phases. However, it was found that the two phases could not be satisfactorily etched by the one etchant, and thus the following two etchants were used :

$\alpha$ Phase	$\beta$ Phase
5 ml. ammonia (0.880 sp. gr.) 5 ml. water 10 ml. 2½% solution of ammonium persulphate	10 g. ferric chloride 1 ml. hydrochloric acid, conc. 500 ml. water

The normal procedure was to etch the specimen for 2-5 seconds in the  $\alpha$  phase etchant, then for about 1-2 seconds in the  $\beta$  phase etchant. The degree of deformation or recrystallization had a marked effect on the etching characteristics of the phases, so no uniform treatment could be employed. Microscopic observations were made and photomicrographs always taken near the portions of the wires from which the X-ray patterns were obtained.

The X-ray diagram of the alloy contained the diffraction patterns of both phases. The grain size of the original material was such that a number of spots appeared on the Debye-Scherrer circles. After deformation the circles were smeared out to continuous rings. At the beginning of recrystallization some



sharp spots which were superimposed on the continuous rings became visible, and their number increased as recrystallization extended through the whole volume. When recrystallization was complete the rings were again spotted.

From the tensile specimens, back reflection diagrams were made with iron K radiation. The very intense background made it necessary, in a few cases, to employ radiation which had been monochromatized with a pentaerythritol crystal.

From the wires, diagrams were obtained in a cylindrical camera (57.3 mm. diameter) using Cu K  $\alpha$  radiation and a pinhole collimating system of 1 mm. diameter and 3 cm. length.

The extent of recrystallization as estimated from the diagrams was classified in the terms—recrystallization just commenced, in progress, and complete. The results so obtained were in good agreement with those of the microscopic examination. This means that the volume irradiated by the X-rays was a representative sample of the specimen.

### III. RESULTS

#### A. THE DEFORMATION OF $\alpha$ - $\beta$ BRASS

The relative deformation of the two phases in  $\alpha$ - $\beta$  brass was followed by microscopic examination of small compression specimens, one free face of which had been electrolytically polished. After approximately 1 per cent. compression, slip lines were visible in the  $\alpha$  phase, while none had appeared in the  $\beta$  phase. After 10 per cent. compression the  $\alpha$  phase showed very heavy slip and surface rumpling, while the  $\beta$  phase was only slightly affected.

In tensile tests the same features were observed. After 27 per cent. elongation, the  $\alpha$  phase showed heavy multiple slip and rumpling whereas the  $\beta$  phase showed only light slip (Plate 1, Fig. 1). The rumpled surface of the specimen was then ground and again electrolytically polished. When further strain was applied to the specimen, it was found that the slip still occurred first in the  $\alpha$  phase. After stressing nearly to fracture (62.5 per cent. elongation), the microscopic examination showed that the  $\beta$  phase was now also heavily deformed although the deformation of the  $\alpha$  phase was still more pronounced (Plate 1, Fig. 2). The specimen was then ground and polished once more. It was stressed to fracture, and micro-examination at the zone of necking showed that again there was heavier slip in the  $\alpha$  phase than in the  $\beta$  phase. However, at the edge of the fracture, the slip lines were very marked in both phases (Plate 1, Fig. 3).

A feature of the deformation is that some slip lines proceed from the  $\alpha$  phase into the  $\beta$  phase with little change in direction (Plate 1, Fig. 4). Since the  $\alpha$  phase has precipitated from the  $\beta$  phase in such a way that one of the (110) planes in the  $\beta$  phase is parallel to one of the (111) planes in the  $\alpha$  phase, and since these planes are the slip planes in their respective structures, the occasional continuity of the slip lines is quite understandable.

## B. RECRYSTALLIZATION OF THE DEFORMED ALLOY

(a) *Experiments with Tensile Specimens*

In the first experiments on the recrystallization of the two phases, tensile deformation was employed as this type of deformation is more uniform than most other deformation processes. The specimens, which had a gauge length of 5 inches, were first annealed for 1 hour at 500° C. and then strained by various amounts. The uniform portions were cut into a number of specimens  $\frac{1}{2}$  inch long, which were annealed for 30 minutes at different temperatures. X-ray back reflection photographs were then made. Table 1 summarizes the results obtained from some of the X-ray photographs.

TABLE 1  
RECRYSTALLIZATION AFTER TENSILE DEFORMATION

Per Cent. Elongation	Phase	State of Recrystallization after Annealing at			
		400° C.	425° C.	450° C.	500° C.
18	$\alpha$	None	Just commenced	Commenced	—
	$\beta$	None	None	None	
24	$\alpha$	Partly	Partly	Fully	Fully
	$\beta$	None	None	Partly	Fully
30	$\alpha$	Partly	Fully	Fully	—
	$\beta$	None	Commenced	Partly	
36	$\alpha$	Partly	Fully	Fully	Fully
	$\beta$	None	Partly	Fully	Fully

The  $\alpha$  phase recrystallizes at lower deformations and temperatures than the  $\beta$  phase, as might be expected from the fact that it is softer and more ductile than the  $\beta$  phase. The state of the  $\alpha$  phase as determined by microscopic examination was in agreement with that deduced from the X-ray photographs. The microscopic examination of the  $\beta$  phase was unconvincing, and it was difficult to decide whether or not any recrystallization had occurred. Very coarse grained tensile specimens were prepared to simplify the micro-examination, but even after 36 per cent. elongation followed by annealing at 600° C., it was found that the original large  $\beta$  grains were practically unchanged. Small  $\beta$  grains appeared only in limited areas. It is possible that the maximum deformation used in the tensile test was insufficient to cause complete recrystallization of the  $\beta$  phase on annealing, and that the increase in sharpness of the diffraction rings was due at least in part to recovery.

(b) *Experiments with Wires*

The experiments using tensile specimens were severely limited in respect to the total deformation that could be applied to the specimen before fracture occurred. Thus, to extend the experiments to cover greater deformation, it was necessary to use deformation processes such as rolling or wire drawing which normally allow a greater deformation to be applied to the specimen. Further, the results of microscopic examination described above indicated that higher deformations would be necessary to make a complete study of the characteristics of the  $\beta$  phase on annealing.

Wire drawing was chosen as the means of deformation and it was found that an  $\alpha$ - $\beta$  brass wire containing up to 80 per cent.  $\beta$  phase could be drawn to nearly 90 per cent. reduction in area. Wire drawing had two other advantages. Firstly, the elongated shape of the grains of the two phases made microscopic examination easier, and secondly, the wire, after etching to suitable size, could be conveniently used to take X-ray transmission photographs. The X-ray back reflection method used for the tensile specimens was not entirely suitable as comparisons had to be made on lines differing markedly in intensity and clarity. With the cylindrical camera, however, the number of lines available was much greater and the background did not make interpretation difficult.

TABLE 2  
RECRYSTALLIZATION AFTER WIRE DRAWING (86% REDUCTION IN AREA)  
EFFECT OF TEMPERATURE

Temperature (° C.)	Degree of Recrystallization	
	$\alpha$ Phase	$\beta$ Phase
300	Commenced	None
350	Partly	None
400	Fully	Commenced
450	Fully	Nearly complete
500	Fully	Fully

(i) *Wires quenched from 750° C.*—Prior to drawing, short lengths of the brass wire were annealed for 15 minutes at 750° C., then water quenched, thus producing a structure which consisted of 80 per cent.  $\beta$  phase and 20 per cent.  $\alpha$  phase. After annealing at different temperatures, the wires were examined by X-rays and microscopically. The states of the two phases in the specimens are shown in Table 2. Plate 2, Figures 5–7, shows typical X-ray photographs of specimens annealed at 300°, 400°, and 450° C.

The  $\alpha$  phase starts to recrystallize at 300° C. but the  $\beta$  phase not until 400° C.

The effect of increasing deformation is shown in Table 3. Even after the heaviest deformation used (86 per cent. reduction in area), the  $\beta$  phase did not recrystallize at either of the two temperatures investigated. As would be expected from experiments on single phase alloys, the  $\alpha$  phase recrystallizes at lower temperatures the higher the deformation.

TABLE 3  
RECRYSTALLIZATION AT 300° AND 350° C. AFTER WIRE DRAWING  
EFFECT OF DEFORMATION

Per Cent. Reduction in Area	Phase	Degree of Recrystallization at	
		300° C.	350° C.
25	$\alpha$	None	Commenced
	$\beta$	None	None
50	$\alpha$	Just commenced	Partly
	$\beta$	None	None
70	$\alpha$	Just commenced	Partly
	$\beta$	None	None
86	$\alpha$	Commenced	Partly
	$\beta$	None	None

Microscopic examination also revealed that the  $\alpha$  phase had commenced to recrystallize at 300° C. and was fully recrystallized at 400° C. However, accurate observation of the  $\beta$  phase was not possible, because the quenched  $\beta$  phase was unstable at the temperatures used for annealing. This resulted in precipitation of the  $\alpha$  phase in the  $\beta$  grains, which had already commenced on annealing at 300° C., but became progressively greater in specimens annealed at 400° and 450° C. (Plate 3, Figs. 8-10). However, the micro-examination did show that the spots on the X-ray diffraction lines from the  $\alpha$  phase were at least partly due to recrystallization (as distinct from precipitation).

(ii) *Stabilized Wires*.—In order to eliminate precipitation of the  $\alpha$  phase which might affect both the X-ray and microscopic examinations, specimens of wire prior to drawing were annealed for 15 minutes at 750° C. and then slowly cooled in the furnace. When the furnace reached the temperature at which the wires were to be annealed *after* drawing, it was maintained at this temperature for 1 week. The specimens were then removed and quenched in water. In this way "stabilized" specimens were obtained in which no redistribution of the quantities of the two phases would occur provided that, after deformation, they were annealed at the temperature used for stabilization prior to drawing. Wires were stabilized at 300° and 350° C., and then subjected to various degrees of



deformation. The X-ray results obtained with specimens stabilized at 300° C. are shown in Table 4.

The microscopic examination confirmed the X-ray results. For example, Plate 4, Figures 11-14, shows the corresponding X-ray photographs and photomicrographs for wires stabilized at 300° and 350° C. and subsequently annealed at these temperatures after 50 per cent. reduction in area by drawing.

TABLE 4  
RECRYSTALLIZATION OF WIRES STABILIZED AT 300° C.

Deformation Per Cent. Reduction in Area	Degree of Recrystallization	
	$\alpha$ Phase	$\beta$ Phase
25	None	Just commenced
50	Commenced	Commenced
70	Partly	Partly
86	Fully	Fully

These results show that the  $\beta$  phase commences to recrystallize *before* the  $\alpha$  phase, and at higher temperatures the recrystallization of both phases occurs simultaneously. Whereas in the previous experiments after 86 per cent. deformation the  $\beta$  phase did not commence to recrystallize until 400° C., in these experiments the  $\beta$  phase was fully recrystallized at 300° C. Experiments with wires stabilized at 350° C. gave similar results.

The alloy used in these experiments had the same composition as that used in the previous experiments. The only difference was that in the former case the alloy was *quenched* from 750° C. prior to drawing, and in the latter, it was slowly cooled to the stabilizing temperature. Thus the alloys may differ in two ways:

- (i) In relative quantities of the two phases; this involves changes in chemical composition.
- (ii) In properties of the  $\beta$  phase; the quenched  $\beta$  phase may differ from the slowly cooled  $\beta$  phase because of the order-disorder change in the vicinity of 460° C.

These factors were investigated in the following experiments.

(c) *The Effect of the Relative Quantities of the Phases*

Short lengths of the brass wire were annealed for 15 minutes at 750° C. and one length was quenched in water. The furnace was then allowed to cool slowly and other lengths of the wire were quenched from 650° and 550° C. In this way, a series of wires was obtained which contained approximately 80, 50, and 30 per cent.  $\beta$  phase respectively. All the wires were then drawn to 86 per cent. reduction in area, and small specimens were annealed at various temperatures. A summary of the X-ray results is shown in Table 5.

These results indicate that the  $\beta$  phase recrystallizes at lower temperatures, the lower the  $\beta$  content of the alloy. For example, after annealing at 400° C., the  $\beta$  phase has only commenced to recrystallize in the 80 per cent.  $\beta$  alloy, is partly recrystallized in the 50 per cent.  $\beta$  alloy, and is fully recrystallized in the 30 per cent.  $\beta$  alloy. The X-ray photographs taken on these three specimens are shown in Plate 5, Figures 15–17; in Plate 5, Figure 15, the  $\beta$  lines are practically continuous, whereas in Plate 5, Figure 17, the lines have broken down to a series of sharp spots indicating complete recrystallization.

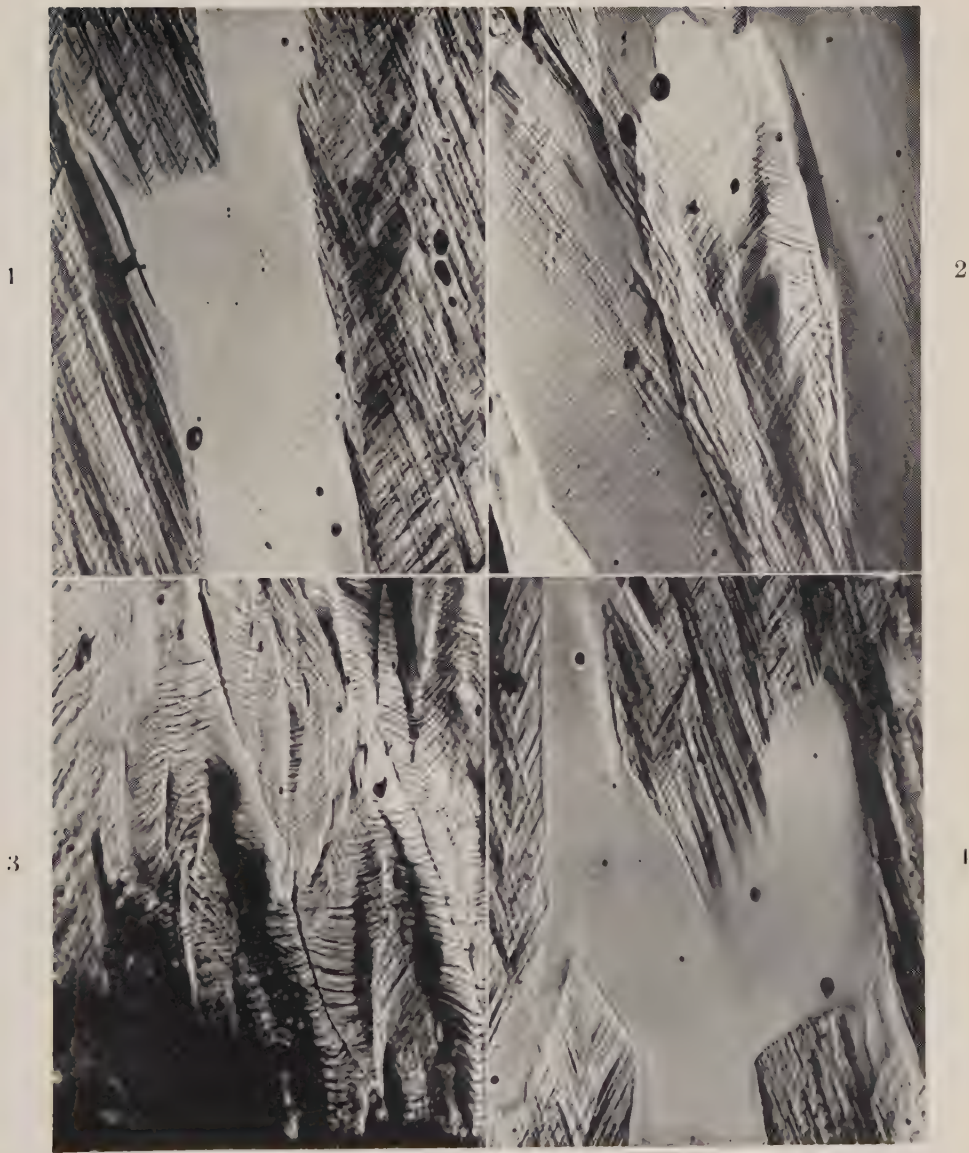
TABLE 5.

RECRYSTALLIZATION OF WIRES QUENCHED FROM DIFFERENT TEMPERATURES

Temperature of Annealing (° C.)	Phase	Specimen Quenched from		
		750° C. (80% $\beta$ )	650° C. (50% $\beta$ )	550° C. (30% $\beta$ )
300	$\alpha$ $\beta$	Partly None	Partly None	Nearly complete None
350	$\alpha$ $\beta$	Partly None	Nearly fully Just commenced	Fully Commenced
400	$\alpha$ $\beta$	Fully Just commenced	Fully Partly	Fully Fully
450	$\alpha$ $\beta$	Fully Nearly complete	Fully Fully	Fully Fully
500	$\alpha$ $\beta$	Fully Fully	Fully Fully	Fully Fully

Over the range of temperature from which the alloys were quenched the solid solubility of zinc in the  $\alpha$  phase *increases* with decreasing temperature, and the zinc content of the  $\beta$  phase in equilibrium with the  $\alpha$  phase also increases with decreasing temperature. The  $\beta$  quenched from 750° C. contains approximately 40 per cent. zinc whereas the  $\beta$  quenched from 550° C. contains 43 per cent. Thus if the two compositions of  $\beta$  were deformed to the same degree, it would be expected that recrystallization would occur at a lower temperature in the  $\beta$  with the lower zinc content, that is, quenched from 750° C. Since in our experiments the reverse was observed, this effect must be small compared with the effect of the varying proportions of the two phases. Thus it appears that the difference in behaviour must be explained in terms of the extent of the deformation of the  $\beta$  phase.

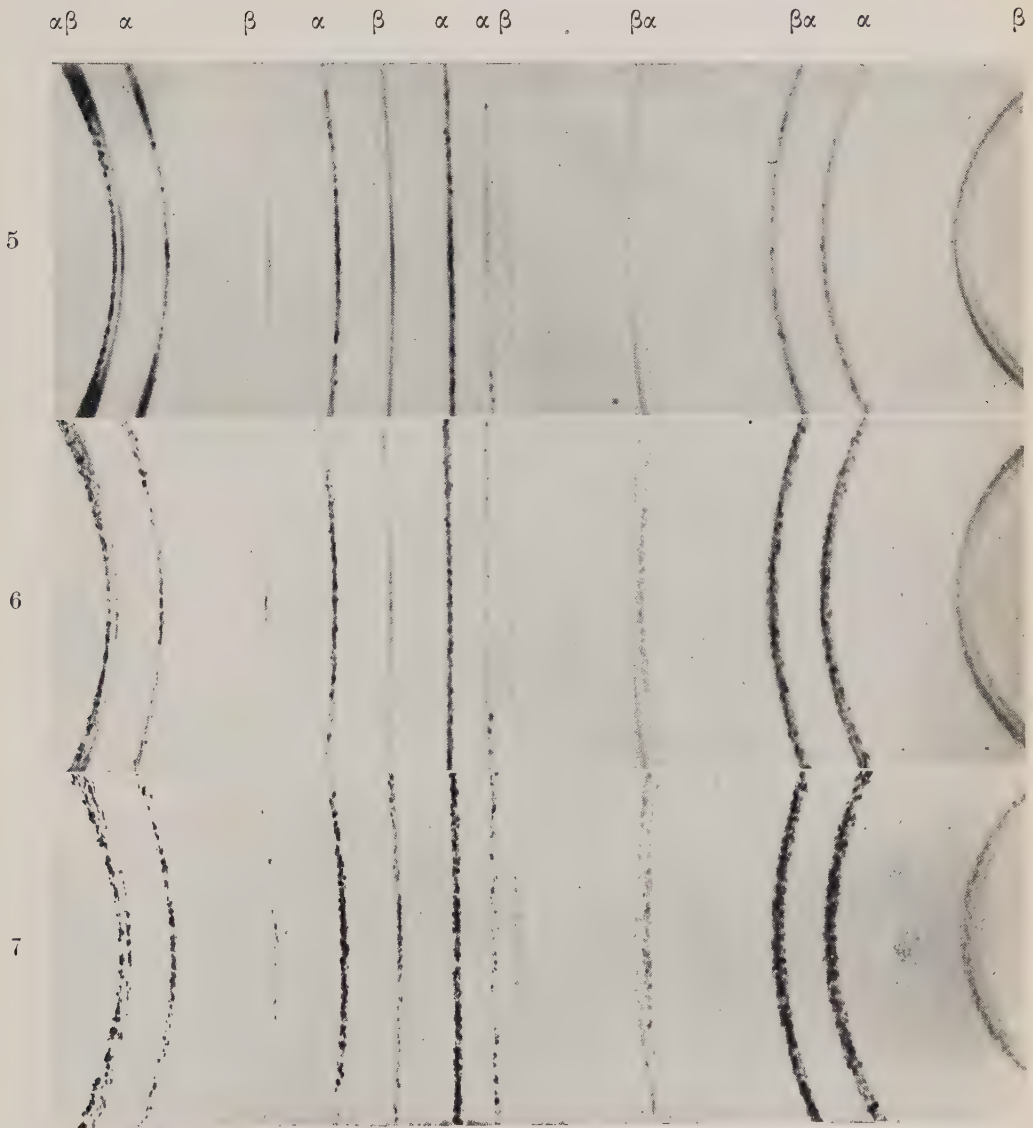
However, in none of the specimens examined did the  $\beta$  phase commence to recrystallize simultaneously with the  $\alpha$  phase. In the specimens quenched from



HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY  
CONTAINING TWO PHASES

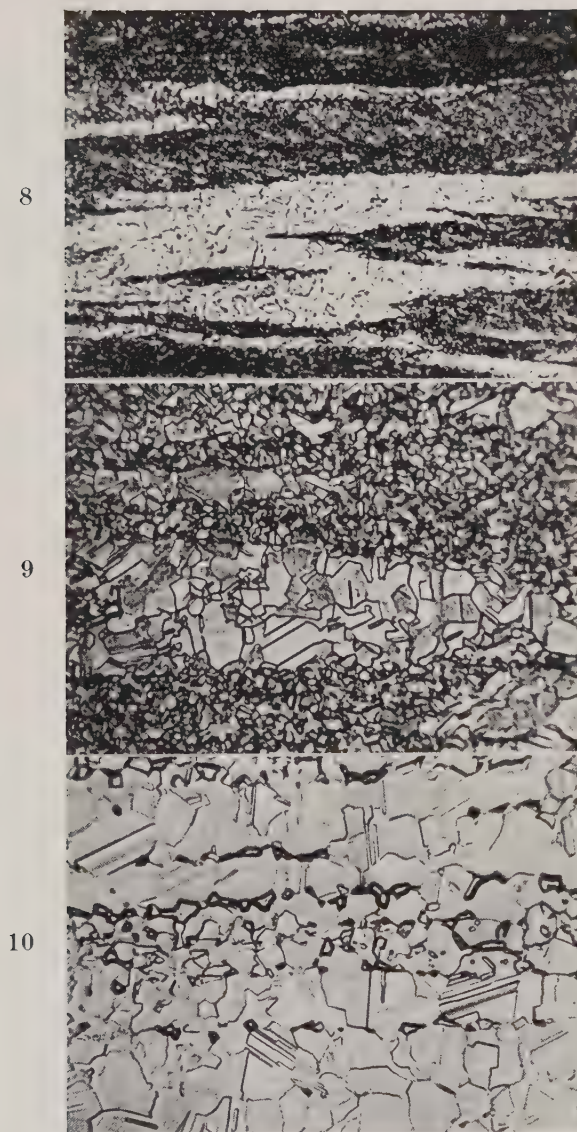






HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY  
CONTAINING TWO PHASES





HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY  
CONTAINING TWO PHASES





$\alpha\beta$     $\alpha$     $\beta$     $\alpha$     $\beta$     $\alpha$     $\alpha\beta$     $\beta\alpha$     $\beta\alpha$     $\alpha$     $\beta$

11

12



13



14



HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY CONTAINING TWO PHASES

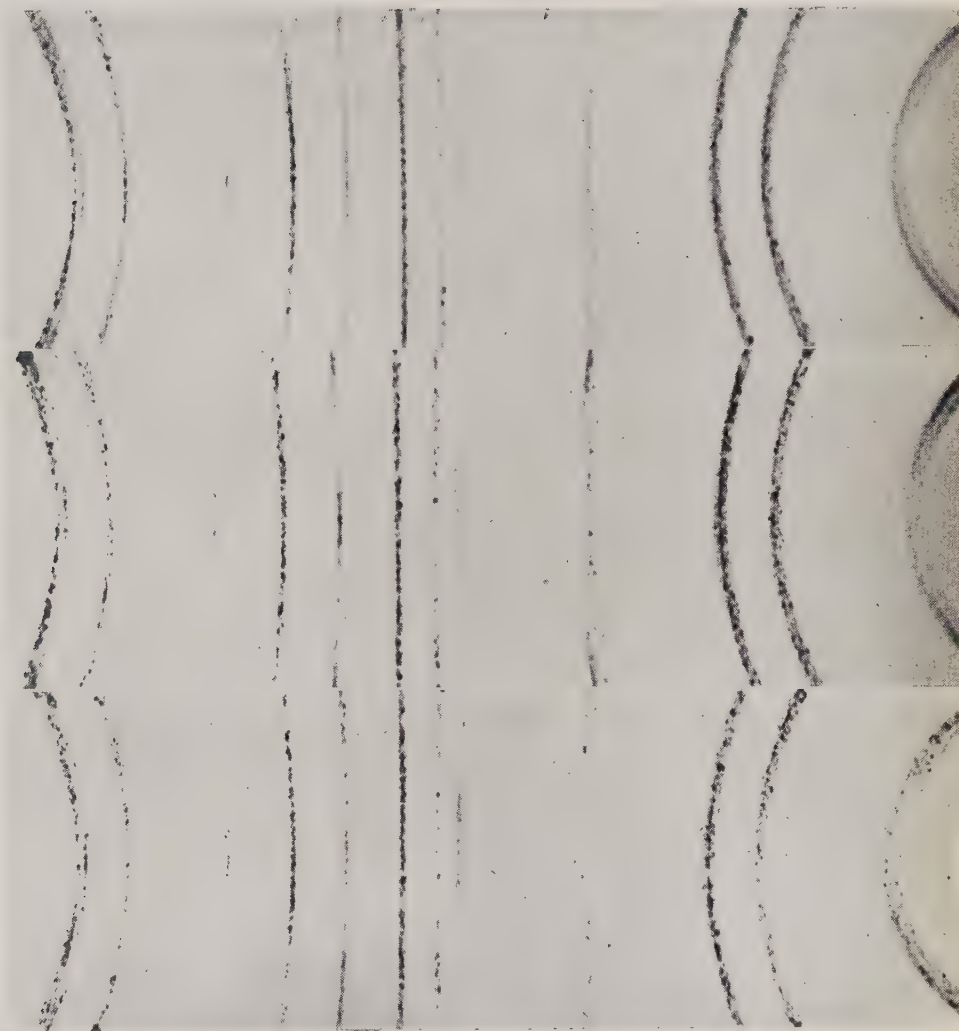


$\alpha\beta$     $\alpha$     $\beta$     $\alpha$     $\beta$     $\alpha$     $\alpha\beta$     $\beta\alpha$     $\beta\alpha$     $\alpha$     $\beta$

15

16

17



HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY  
CONTAINING TWO PHASES





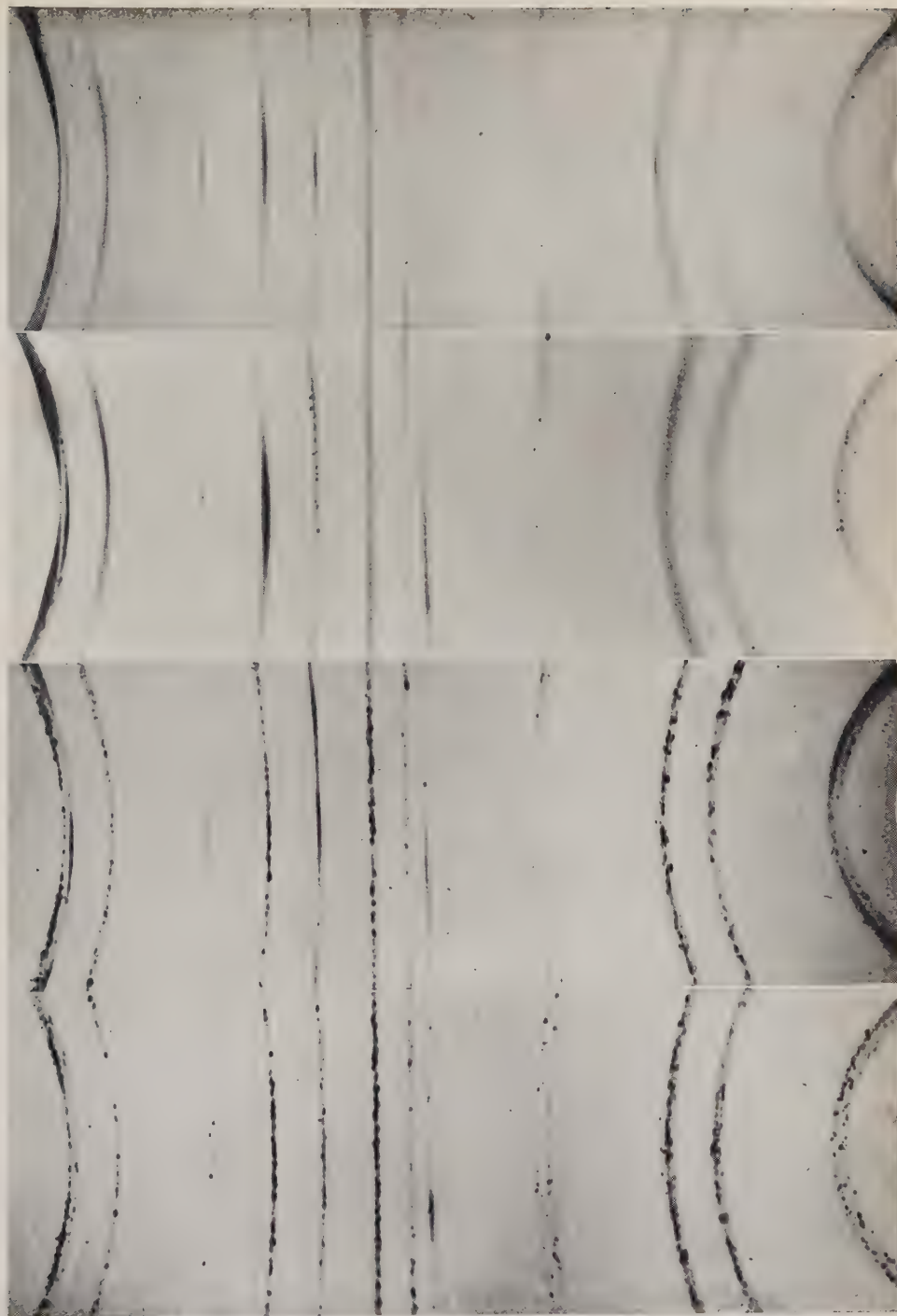
$\alpha\beta$   $\alpha$   $\beta$   $\alpha$   $\beta$   $\alpha$   $\alpha\beta$   $\beta\alpha$   $\beta\alpha$   $\alpha$   $\beta$

18

19

20

21







HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY  
CONTAINING TWO PHASES





28



29



30

HONEYCOMBE and BOAS.—THE DEFORMATION AND RECRYSTALLIZATION OF AN ALLOY  
CONTAINING TWO PHASES



750° C., the  $\beta$  phase commenced to recrystallize 100° C. above the  $\alpha$  phase, while this difference was still about 50° C. in the alloy containing 30 per cent.  $\beta$ . Thus it appeared that the rate of cooling below 550° C. before deformation was responsible for the marked change in the behaviour of the  $\beta$  phase on recrystallization. This suggested that an effect associated with the order-disorder transformation rather than redistribution of the phases causes the  $\beta$  to nucleate before the  $\alpha$  phase.

(d) *The Effect of the Order-Disorder Transformation*

Two lengths of wire were annealed for 15 minutes at 800° C. and were then slowly cooled in the furnace to 475° C. One specimen was then water quenched, while the other remained in the furnace till it reached a temperature of 400° C. After being held at 400° C. for 30 minutes the second piece of wire was quenched in water. Both wires were drawn to 86 per cent. reduction in area immediately after quenching. As in the previous experiments, small lengths of the drawn wire were annealed at various temperatures, then examined both microscopically and by X-rays. The results of the X-ray examination are shown in Table 6.

TABLE 6  
RECRYSTALLIZATION OF WIRES QUENCHED FROM ABOVE AND BELOW THE CRITICAL TEMPERATURE

Annealing Temperature (° C.)	Phase	Degree of Recrystallization	
		Wire Quenched from 475° C.	Wire Quenched from 400° C.
250	$\alpha$	None	None
	$\beta$	None	Commenced
275	$\alpha$	Commenced	Partly
	$\beta$	None	Partly
300	$\alpha$	Almost fully	Almost fully
	$\beta$	Commenced	Almost fully
350	$\alpha$	Fully	Fully
	$\beta$	Partly	Fully

In the wire originally quenched from *above* the critical temperature (454° C.), the  $\beta$  phase started to recrystallize at a *higher* temperature than the  $\alpha$  phase, whereas in the wire originally quenched from *below* the critical temperature, the  $\beta$  phase commenced to recrystallize at a *lower* temperature than the  $\alpha$  phase. The inversion of the order of recrystallization of the two phases is illustrated in Plate 6, Figures 18-21, which shows the X-ray photographs of the two series of

specimens annealed at 250° and 300° C. Plate 7, Figures 22-27, shows the correlation between the results of X-ray examination and those of microscopic examination. The commencement of recrystallization in the  $\beta$  phase before the  $\alpha$  phase is shown in Plate 7, Figure 23.

This inversion of the order of recrystallization cannot be due to a change in solid solubility, as the phase boundaries of the  $(\alpha + \beta)$  region are nearly vertical in the temperature range concerned (400°-500° C.). For the same reason no change in the proportions of the two phases occurs. It is therefore concluded that the rate of cooling through the range of the critical temperature determines the order of recrystallization.

#### *(e) The Mode of Recrystallization of the Phases*

The microscopic method of examination has one advantage over the X-ray method, namely, it reveals where the recrystallization commences in the grains of the two phases. The recrystallized grains of each phase can be detected and readily distinguished, as only the crystals of the face-centred cubic  $\alpha$  phase show annealing twins.

Examination of specimens in which recrystallization had just commenced revealed that the  $\beta$  phase tended to nucleate at the boundary with the  $\alpha$  phase. This occurred independently of whether the  $\beta$  phase commenced to recrystallize at a higher or lower temperature than the  $\alpha$  phase (Plate 8, Figs. 28 and 29). On the other hand, the  $\alpha$  phase did not appear to nucleate preferentially at the boundaries of the two phases. Nucleation occurred more at random throughout the crystals of the phase (Plate 8, Fig. 29).

### IV. DISCUSSION AND CONCLUSIONS

Our experiments on the deformation of duplex brass both by compression and tension show clearly that the two phases do not start deforming simultaneously. The  $\alpha$  phase starts to deform plastically under lower stress than the  $\beta$  phase, even in a very slowly cooled alloy in which the  $\beta$  phase would recrystallize at a lower temperature than the  $\alpha$  phase. In the later stages of the tensile test the  $\alpha$  phase still appears to deform more severely than the  $\beta$  phase. Clearly the work hardening curves of the two phases do not intersect, but this behaviour is not necessarily a general feature of duplex alloys.

These conclusions are based on observations of slip lines and rumpling on the surfaces of the specimens. It may be asked whether the observations give a true indication of the relative deformation of the two phases. Cases are known even in single crystals where deformation occurs without the formation of visible slip lines. However, in our tensile specimen, the  $\beta$  grains at the point of fracture showed as heavy signs of deformation as the  $\alpha$  grains, and for this reason we believe that the visual evidence truly represents the relative deformation of the two phases.

The experiments on recrystallization show that the two phases do not recrystallize simultaneously. Whether the  $\alpha$  or the  $\beta$  phase recrystallizes first depends on the treatment prior to deformation, and it seems that the phase in



which the recrystallization commences is determined by the rate of cooling through the range near the critical temperature of the order-disorder transformation. The disordered state of  $\beta$  brass cannot be preserved at room temperature by quenching. However, Smith(3) has found that the hardness of quenched  $\beta$  brass is higher than that of slowly cooled  $\beta$  brass. Correspondingly the work hardening characteristics of quenched and slowly cooled  $\beta$  brass may be different, particularly during the early stages of deformation. In the later stages of deformation the ordered  $\beta$  structure probably becomes disordered and the difference between the two states of the  $\beta$  phase would thus be eliminated. It is suggested that the different behaviour of the wires on recrystallization is due to different work hardening characteristics in the *initial* stages of deformation.

The experiments with wires containing different amounts of the  $\beta$  phase showed that this phase recrystallizes at lower temperatures the smaller the proportion in the alloy. This means that the deformation per unit volume of the  $\beta$  phase is larger the smaller the quantity of  $\beta$  phase embedded in the  $\alpha$  phase. As has been stated above, the two phases do not take part in the deformation to equal extents. For a given total deformation of the duplex alloy, the deformation per unit volume of each phase must vary with the relative proportions of the two phases present. The experiments show that the deformation of the  $\beta$  phase can be carried to a larger extent without failure, the higher the proportion of  $\alpha$  phase present. A pure  $\beta$  brass is relatively brittle and can only be given about 20 per cent. elongation in tension before fracture, whereas an  $\alpha$ - $\beta$  brass can be given at least 40 per cent. elongation, and this difference in behaviour becomes still more marked if wire drawing is employed as a means of deformation.

From Table 5 it is clear that the  $\alpha$  phase as well as the  $\beta$  phase tends to recrystallize more readily the greater the amount of  $\alpha$  phase present, although the effect here is not as pronounced. This difference can possibly be attributed to the greater probability of finding recrystallized  $\alpha$  grains if the volume irradiated by the X-ray beam contains a large proportion of  $\alpha$  than if that volume contains a small proportion of  $\alpha$ . On the other hand, the  $\beta$  phase appears, from X-ray photographs, to recrystallize more readily the less the amount of  $\beta$  present. Hence the true effect in this case must be still more pronounced than is apparent from inspection of X-ray photographs.

From the above evidence, it is clear that the  $\alpha$  phase assists the  $\beta$  phase deform, and thus some interaction must occur between the two phases. It would therefore be expected that the  $\beta$  phase bordering the  $\alpha$  phase would be more deformed than the interior of the  $\beta$  crystals, and correspondingly that the deformation of the  $\alpha$  phase would be restricted in the vicinity of  $\beta$  grains. Evidence for the former is that the slip lines in the  $\beta$  phase are often stronger near the boundary with the  $\alpha$  phase. This is further confirmed by the behaviour of the two phases on recrystallization. The  $\beta$  grains usually nucleate at or near the  $\alpha$ - $\beta$  boundaries while the  $\alpha$  crystals nucleate rather in the interior of the grains.

There is no simple correlation between the relative amounts of deformation of the two phases and their behaviour on recrystallization. The tendency to recrystallize also depends on the melting point, the state of order in the lattice,

and the chemical composition. Although the chemical compositions of the two phases in duplex brass are very similar and the  $\alpha$  phase always deforms to a larger extent, the tendency to recrystallize is not always greater in the  $\alpha$  phase. Thus it is not possible to conclude that a phase has suffered more severe deformation because it recrystallizes at a lower temperature.

## V. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Section of Tribophysics, C.S.I.R.

We wish to thank Dr. S. H. Bastow, Officer-in-Charge, Section of Tribophysics, C.S.I.R., for his interest in the work and for helpful discussions. We are grateful to Professor E. J. Hartung for laboratory facilities in the Chemistry School, University of Melbourne.

## VI. REFERENCES

- (1) UNCKEL, H.—*J. Inst. Met.* **61**: 171 (1937).
- (2) RHINES, F. N., and WARD, R.—*Ibid.* **61**: 193 (1937) (Discussion).
- (3) SMITH, C. S.—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **152**: 144 (1943).

## EXPLANATION OF PLATES 1-8

### PLATE 1

- Fig. 1.—Polished surface of  $\alpha$ - $\beta$  brass after 27 per cent. elongation. The  $\alpha$  phase shows marked slip whereas there is little slip in the  $\beta$  phase.
- Fig. 2.— $\alpha$ - $\beta$  brass after 62.5 per cent. elongation.  $\alpha$  and  $\beta$  both show marked signs of deformation but the  $\alpha$  phase still appears to be more heavily deformed.
- Fig. 3.—Area at edge of fracture.  $\alpha$  and  $\beta$  show marked slip.
- Fig. 4.—Specimen after 25 per cent. elongation. Some of the slip lines continue from the  $\alpha$  to the  $\beta$  phase without change in direction.
- (Photomicrographs.  $\times 200$ .)

### PLATE 2

X-ray powder photographs of  $\alpha$ - $\beta$  brass wires deformed to 86 per cent. reduction in area, then annealed at different temperatures. The wires were quenched from 800° C. prior to drawing.

- Fig. 5.—Annealed  $\frac{1}{2}$  hr. at 300° C.  $\alpha$  commenced to recrystallize,  $\beta$  none.
- Fig. 6.—Annealed  $\frac{1}{2}$  hr. at 400° C.  $\alpha$  fully recrystallized,  $\beta$  commenced.
- Fig. 7.—Annealed  $\frac{1}{2}$  hr. at 450° C.  $\alpha$  fully recrystallized,  $\beta$  nearly complete.
- (Photographs are one and a half times the size of the originals.)

### PLATE 3

Photomicrographs corresponding to X-ray photographs (Plate 2, Figs. 5-7).

- Fig. 8.—Annealed  $\frac{1}{2}$  hr. at 300° C.  $\alpha$  commenced to recrystallize,  $\beta$  still deformed.
- Fig. 9.—Annealed  $\frac{1}{2}$  hr. at 400° C.  $\alpha$  fully recrystallized. The onset of precipitation of the  $\alpha$  and  $\beta$  phases prevents the detection of recrystallization in this phase.
- Fig. 10.—Annealed  $\frac{1}{2}$  hr. at 450° C.  $\alpha$  and  $\beta$  appear to be completely recrystallized.
- (Photomicrographs.  $\times 550$ .)

## PLATE 4

Corresponding X-ray powder photographs and photomicrographs of brass specimens stabilized and annealed at 300° and 350° C. after 50 per cent. reduction in area by drawing.

Figs. 11 and 13.—Stabilized and annealed at 300° C. The  $\alpha$  and  $\beta$  phases have both commenced to recrystallize. The recrystallization in the  $\beta$  phase is somewhat more advanced than in the  $\alpha$  phase.

Figs. 12 and 14.—Stabilized and annealed at 350° C. The  $\alpha$  and  $\beta$  phases are completely recrystallized.

(Photographs are one and a half times the size of the originals. Photomicrographs  $\times 420$ .)

## PLATE 5

X-ray powder photographs of brass wires quenched from various temperatures, drawn to 86 per cent. reduction in area, then annealed for 30 minutes at 400° C.

Fig. 15.—Eighty per cent.  $\beta$  in alloy. The  $\alpha$  phase is fully recrystallized whereas the  $\beta$  phase has just started to recrystallize.

Fig. 16.—Fifty per cent.  $\beta$  in alloy. The  $\alpha$  phase is fully recrystallized and the  $\beta$  phase is partly recrystallized.

Fig. 17.—Thirty per cent.  $\beta$  in alloy. The  $\alpha$  and  $\beta$  phases are both fully recrystallized.

(Photographs are one and a half times the size of the originals.)

## PLATE 6

X-ray photographs of specimens quenched from above and below the critical temperature for order, drawn 86 per cent. reduction in area, then annealed at 250° and 300° C.

Fig. 18.—Quenched from 475° C. Annealed at 250° C.  $\alpha$  and  $\beta$  still deformed.

Fig. 19.—Quenched from 400° C. Annealed at 250° C.  $\alpha$  deformed,  $\beta$  commenced to recrystallize.

Fig. 20.—Quenched from 475° C. Annealed at 300° C.  $\alpha$  almost fully recrystallized,  $\beta$  commenced.

Fig. 21.—Quenched from 400° C. Annealed at 300° C.  $\alpha$  and  $\beta$  almost fully recrystallized.

(Photographs are one and a half times the size of the originals.)

## PLATE 7

Photomicrographs of specimens quenched from above and below the critical temperature for order, drawn 86 per cent. reduction in area, then annealed at different temperatures.

Fig. 22.—Quenched from 475° C. Annealed at 250° C.  $\alpha$  and  $\beta$  still deformed.

Fig. 23.—Quenched from 400° C. Annealed at 250° C.  $\alpha$  still deformed,  $\beta$  commenced to recrystallize.

Fig. 24.—Quenched from 475° C. Annealed at 300° C.  $\alpha$  partly recrystallized,  $\beta$  just commenced.

Fig. 25.—Quenched from 400° C. Annealed at 300° C.  $\alpha$  and  $\beta$  nearly fully recrystallized.

Fig. 26.—Quenched from 475° C. Annealed at 350° C.  $\alpha$  fully recrystallized,  $\beta$  partly recrystallized.

Fig. 27.—Quenched from 400° C. Annealed at 350° C.  $\alpha$  and  $\beta$  fully recrystallized.

(Photomicrographs.  $\times 550$ .)

## PLATE 8

The mode of recrystallization of the phases.

Fig. 28.—The beginning of recrystallization in quenched  $\beta$  phase. Nucleation near the boundary of the two phases.

Fig. 29.—The beginning of recrystallization in slowly cooled  $\beta$  phase. Again nucleation occurs at the boundary.

Fig. 30.—Nucleation of the  $\alpha$  phase occurs at random.

(Photomicrographs.  $\times 550$ .)

---



# LUBRICATION OF METAL SURFACES BY SILICONE FILMS

By J. N. GREGORY\* and MARJORIE J. NEWING\*

(Plate 1)

[Accepted for Publication November 26, 1947]

## Summary

Methods for the preparation of mono-, di-, and tri-isoamyl chlorsilanes, and the monocetyl chlorsilane have been developed. Molecular layers of the mono- and di-ethyl and mono- and di-isoamyl silicones on various metal surfaces have been tested for their boundary lubricating properties. The layers have been shown to be excellent lubricants, with unusual thermal stability and resistance to wear.

## I. INTRODUCTION

The term "silicone" has been applied to that broad group of compounds formed by the controlled hydrolysis of halosilanes and alkoxy silanes. According to the degree and conditions of hydrolysis the products may vary in consistency

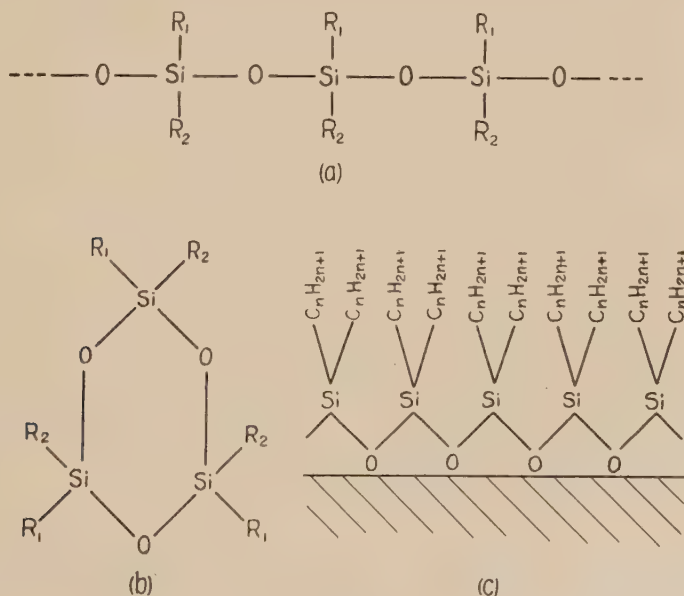


Fig. 1

from transparent oily liquids to hard resinous solids. The hydrolysis of a dialkyl dichlorsilane  $\text{R}_1\text{R}_2\text{SiCl}_2$  may result in the formation of long chains (Fig. 1(a)), or ring structures (Fig. 1(b)), i.e.  $(\text{R}_2\text{SiCl}_2)_n + n\text{H}_2\text{O} \rightarrow (\text{R}_2\text{SiO})_n + 2n\text{HCl}$ ,

\* Section of Tribophysics, C.S.I.R.

while monoalkyl trichlorsilanes form very complex ring structures or three-dimensional networks.

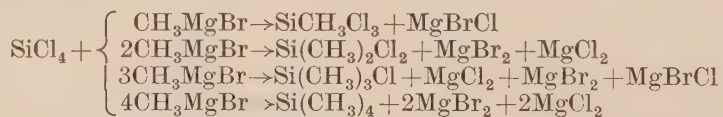
A surface containing adsorbed water exposed to halosilane vapour becomes coated with a tenacious hydrophobic molecular layer of silicone in which the organic radicles are considered(1) to be oriented vertically to the surface (Fig. 1 (c)). When these organic radicles are long chain aliphatic, such a system shows a marked similarity to systems well known for good boundary lubrication, viz. those in which the active heads of long chain polar molecules are attached to a metal surface with the consequent formation of oriented monolayers. Furthermore, in a silicone coating, all the long chains have their ends rigidly connected to each other by means of the -Si-O-Si- chains as shown in the diagram (Fig. 1 (c)). For these reasons, and on account of the great thermal stability of these compounds, it was thought that metal surfaces coated with silicones from relatively long chain alkyl halosilanes might prove useful in boundary lubrication. This paper describes the preparation of several such silanes and the investigation of boundary lubricating properties of silicone layers produced from them.

## II. EXPERIMENTAL

### (A) CHEMICAL INVESTIGATIONS

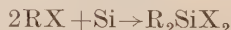
#### (a) *Existing Methods for the Preparation of Halosilanes*

The most general method of preparation is that of Kipping and his co-workers(2, 3, 4, 5) used by Rochow and Gilliam(6) and Hyde and De Long(7). This method makes use of the reaction of a Grignard reagent with silicon tetrachloride. The Grignards were usually the methyl or ethyl magnesium bromide :



The mixture of methyl chlorsilanes can be separated by fractional distillation.

Another method of preparation was devised by Rochow(8). He found it possible to synthesize alkyl and aryl silicon halides by the direct reaction of the organic halide in the gaseous phase at high temperatures with elementary silicon. An increased yield was obtained by the use of a reduced copper catalyst. Under favourable conditions the principal reaction is



but varying amounts of  $\text{R}_3\text{SiX}$ ,  $\text{RSiX}_3$ ,  $\text{SiX}_4$ , and  $\text{R}_4\text{Si}$  are also obtained according to the conditions of the experiment. Both methods, however, only refer to the synthesis of ethyl and methyl compounds.

#### (b) *The Gaseous Phase Reaction in the Preparation of Long Chain Silanes*

In view of the simplicity of this method, an attempt was made to form the three isoamyl silanes by passing the vapour of isoamyl chloride over silicon containing reduced copper at temperatures ranging from 300°–400° C. The issuing vapours were condensed and the liquid fractionally distilled using a

semi-micro fractionating apparatus(9). Table 1 shows the results obtained. In each case 10 ml. of the distillate was used.

TABLE 1

Reaction Temperature (° C.)	Fractions Collected (ml.)			
	33°-40° C.	40°-92° C.	92°-97° C.	> 97° C.
300	0	0.8	9.0	0
320	2.7	1.0	5.5	0.5
330-340	2.7	1.0	5.5	0.5
360	4.5	1.5	3.0	1.0
350-400	3.0	3.4	2.4	1.0

The fraction boiling 92°-97° C. is unchanged isoamyl chloride and that boiling 40°-92° C. mainly silicon tetrachloride. The first fraction (33°-40° C.) is a mobile fuming colourless liquid which gives a white gelatinous precipitate with water. This precipitate dissolves in sodium hydroxide with the evolution of an inflammable gas which explodes violently on sparking with oxygen. The liquid is probably silicochloroform  $\text{SiHCl}_3$  (b.p. 35° C.) which can be made by passing hydrogen chloride gas over silicon with copper catalyst at 350° C. Hydrogen chloride could be present if the elevated temperatures had caused pyrolysis of the isoamyl chloride.

It is obvious from the table that a temperature of 320° C. is not high enough to cause appreciable reaction between the chloride and silicon, but raising the temperature causes pyrolysis of the chloride with the subsequent formation of greater amounts of silicochloroform.

The chlorsilane is contained in that very small amount of mobile yellow fuming liquid which boiled above 97° C. It was identified by its ability to form a water-repellent film on a glass surface. It seems, therefore, that the gaseous phase reaction may not be suitable with long chain halides owing to the predominance of the pyrolysis reaction at the temperatures requires to produce chlorsilanes.

#### (c) *The Grignard Reaction for Long Chain Silanes*

This is the classical method of Kipping(10) in which a Grignard compound of the required alkyl (or aryl) group reacts with a silicon compound, e.g. silicon tetrachloride, forming the chlorsilane and  $\text{MgCl}_2$ . The reaction is carried out in anhydrous ether, diethyl and dibutyl ethers being most commonly used. With the short chain compounds the Grignard ( $\text{RMgX}$ ) is added to a cooled ethereal solution of silicon tetrachloride(11). However, in some of the cases considered

in this work it was found necessary to reflux the reaction mixture for some time before appreciable reaction occurred.

(i) *Isoamyl Silanes*.—It was found possible to prepare the three isoamyl silanes by this method using isoamyl magnesium bromide and silicon tetrachloride, and the yields are summarized in Table 2. In the first column is shown the ratio of the amounts of silicon tetrachloride and Grignard added in an attempt to make the mono-, di-, and tri-isoamyl chlorsilanes. The percentage yields are based on the amount of the added Grignard used in forming the various products collected.

TABLE 2  
YIELDS OF ISOAMYL SILANES

Ratio of SiCl <sub>4</sub> : Grignard	Percentage Yields (based on Grignard)		
	Mono- B.P. 158°–162° C.	Di- B.P. 144°–148° C./53 mm.	Tri- B.P. 180°–184° C./53 mm.
1 : 1	50	—	—
1.5 : 1	63	—	—
1 : 2	12½	14	4
1 : 3	5	7	11

The compounds of constant boiling point were identified by analysis for silicon and chlorine after fractionation. The details of the method of preparation are given in Appendix I.

For analysis of the silanes the method of Gilliam, Liebhafsky, and Winslow(11) for determining silicon and chlorine by means of a Parr bomb fusion was abandoned for the trichlorsilane owing to the violence of its reaction with sodium peroxide. The halogen was successfully determined by hydrolysis of the silane with sodium hydroxide solution, and titration by Mohr's method after neutralization with nitric acid. Numerous methods were tried for the determination of silicon in the mono- and di-isoamyl silanes. The most successful method for the oxidation of the organic radicle was found to be by means of concentrated sulphuric acid in the presence of copper sulphate or a drop of mercury as a catalyst as in the Kjeldahl method for the determination of nitrogen. After oxidation of all the organic matter, the remaining silica was filtered off. The silica adhering to the vessel walls was removed by boiling with caustic soda and the silica recovered by repeated evaporation as in the general method of analysis of soluble silicates(12). The reaction between sodium peroxide and tri-isoamyl monochlorsilane was not violent, and it was possible to use the Parr bomb fusion method for the silicon determination in this case.

The main portion of the yield from the mono-isoamyl silane experiment, after removal of ether and silicon tetrachloride, boiled at 158°–162° C./760 mm.



This gave a chlorine content of 44.57 per cent. The theoretical value is 51.7 per cent. As the chlorine content was reproducible the discrepancy was thought to arise either from partial substitution of the chlorine attached to the silicon by bromine from the isoamyl magnesium bromide, or from the organic radicle if it contained some higher homologues. A potentiometric titration with silver nitrate, using a silver electrode, showed that the amount of bromide present was only sufficient to give a true halogen content of 46.2 per cent. To eliminate the effect of the presence of higher homologues on the results of the analysis the criterion for the silanes was taken to be the silicon to chlorine ratio.

When, in attempts to prepare the di- and tri-isoamyl silanes, the quantity of Grignard reagent exceeded that required to form the mono-isoamyl compound, a proportion of mono-isoamyl silane was always produced. However, unlike the product of equimolecular proportions of silicon tetrachloride and the Grignard reagent, which gave an almost theoretical silicon and chlorine ratio for mono-isoamyl trichlorsilane, the analysis of this compound shows very low and variable chlorine and silicon contents.

In both cases the fraction boiling from 158°–162° C. was taken to be the mono-isoamyl silane and it is now considered that the discrepancy is due to the formation of the hydrocarbon  $C_{10}H_{22}$  (2, 7-dimethyloctane) by a reaction similar to that occurring in the cetyl Grignard reaction (see (c) (ii) below). The reaction to form the di- and tri-isoamyl compounds is slow compared with that forming the mono-compound, and the consequent prolonged refluxing (see Appendix I) of the excess Grignard probably favours the formation of this hydrocarbon, which has a boiling point of 159.6° C. and would thus be included in the mono-isoamyl fraction.

In an attempt to prevent the formation of this hydrocarbon, 1 mole of Grignard in ether was added very slowly to 1.5 moles of silicon tetrachloride in ether. This gave a yield of 63 per cent. of mono-isoamyl trichlorsilane with a chlorine content of 47.5 per cent. The silicon content was 12.5 per cent., which gives a silicon to chlorine ratio of 1 : 3.01. A further attempt to prevent the formation of the paraffin, by adding the magnesium slowly to the mixture of silicon tetrachloride and isoamyl bromide in ether (thus making the Grignard *in situ*), failed owing to the reaction between silicon tetrachloride and magnesium which made it necessary to add quite a large excess of magnesium. Some of this excess magnesium probably reacted with the isoamyl bromide to form the paraffin, as the chlorine content obtained was only about 30 per cent.

In the di-isoamyl dichlorsilane preparation the results obtained were as follows :

TABLE 3

Samples	I	II	III	IV	V
B.P. of fraction (° C.) ..	90–104/ 124 mm.	104–120/ 124 mm.	140–150/ 50 mm.	150–162/ 50 mm.	170–210/ 50 mm.
Chlorine content (%) ..	37	37	26.5	22	12

From Table 3 it seems probable that sample III contains most of the di-isoamyl silane (theoretical halogen content 34 per cent.), as there seems to be a distinct break in the boiling point range before this point. Redistillation of this sample gave a product boiling  $144^{\circ}$ – $148^{\circ}$  C./56 mm. with a chlorine content of 26.2 per cent. and a silicon content of 10.7 per cent.—the silicon to chlorine ratio being 1 : 1.94. Sample V, of which a small amount only was obtained, may contain the tri-isoamyl silane (theoretical chlorine content 12.8 per cent.).

In the tri-isoamyl preparation the mono- and di-isoamyl silane samples were collected and the boiling points of the fractions are shown in Table 4.

TABLE 4

Samples	I	II	III
B.P. ( $^{\circ}$ C.) .. ..	152–158/ 55 mm.	160–170/ 55 mm.	178–190/ 55 mm.
Chlorine content (%)	26	16.5	11

Here it seems that sample III contains the tri-isoamyl silane. Redistillation gave a sample boiling  $184^{\circ}$ – $188^{\circ}$  C./55 mm. with a halogen content of 13.2 per cent. and a silicon content of about 10 per cent., giving a silicon to chlorine ratio of approximately 1 : 1. The fact that the silicon to chlorine ratio is not exactly 1 : 1 may be due to incomplete fractionation, as the amount of tri-isoamyl silane formed in any experiment is small, and also to the difficulty of getting consistent silicon analysis results. A summary of the analyses of these compounds is given in Table 5.

TABLE 5

Formula	Chlorine (%)	Silicon (%)	Cl : Si
(C <sub>5</sub> H <sub>11</sub> )SiCl <sub>3</sub> ..	47.5	12.5	3.01 : 1
(C <sub>5</sub> H <sub>11</sub> ) <sub>2</sub> SiCl <sub>2</sub> ..	26.2	10.7	1.94 : 1
(C <sub>5</sub> H <sub>11</sub> ) <sub>3</sub> SiCl ..	13.2	c. 10	c. 1 : 1

(ii) *Monocetyl trichlorosilane*.—It was also found possible to prepare monocetyl trichlorosilane (b.p.  $180^{\circ}$ – $190^{\circ}$  C./5 mm.) by the Grignard method. The yield, however, was very small owing to the occurrence of side reactions. A large quantity of dicetyl (C<sub>32</sub>H<sub>66</sub>) precipitated from the ether solution. This reaction, according to Oldham and Ubbelohde(13), takes place when long chain aliphatic halides react with magnesium in absolute ether in a type of Wurtz reaction, i.e.  $2\text{C}_{16}\text{H}_{33}\text{I} + \text{Mg} \rightarrow \text{C}_{32}\text{H}_{66} + \text{MgI}_2$ .

The commercial cetyl alcohol used in this preparation, although it was distilled, had a range of boiling point of about 15° C. Thus it was not possible to purify further by distillation the monocetyl silane (b.p. 180°–190° C./5 mm.).

As the product of hydrolysis of the cetyl silane is a white waxy solid, it is possible, in this case, merely to hydrolyse a weighed amount of the silane and then filter. The filtrate can be titrated with standard silver nitrate after neutralization to determine the chlorine content. In this way the chlorine content was found to be 28.3 per cent. and the silicon 7.7 per cent., giving a silicon to chlorine ratio of 1 : 2.8. The silicon was determined by ignition of the solid hydrolysis product.

#### (B) FRICTIONAL PROPERTIES OF SILICONE FILMS ON METAL SURFACES

An oriented monolayer of a silicone on a metal surface, as shown in a previous diagram (Fig. 1 (c)), would be expected to be water repellent because the face presented is essentially hydrocarbon in nature. Measurements of the contact angles of the system air—water—siliconed surface by the Taggart bubble method, for several different surfaces, show that even the ethyl compound gives a contact angle of 85°–90°, and increase in chain length does not appear to have any significant effect on this property. Contact angles were taken, therefore, as an indication of the uniformity and completeness of a silicone layer on any metal.

The metal surfaces were metallurgically polished to a 600 paper, electrolytically cleaned and then dried in air after immersion in hot distilled water. They were then exposed to a humid atmosphere at 35° C. for about half an hour prior to the silicone treatment to ensure that there were sufficient adsorbed water molecules on the surfaces to react with the silane vapour. Dried air was bubbled through the various liquid silanes and then passed over the surfaces. The length of time necessary to obtain an even coating (as tested by contact angle and subsequent friction measurements) varied with the volatility of the silane. Mono- and di-ethyl silanes\* required treatment of about one hour while the mono-isoamyl silane seemed to need about twice this time. The di-isoamyl silane (b.p. 144°–148° C./56 mm.) did not appear to give uniform coating even after four hours of exposure to the silane vapour. This is probably due to the much higher boiling point of this silane. The difficulty was overcome by carrying out the treatment at a reduced pressure (c. 52 mm. of Hg) whilst heating the silane liquid on a water-bath. After four hours a satisfactory coating was obtained. Different metals with the same silane coatings all showed similar contact angles.

On two of the metal surfaces (viz. copper and steel) a film was often clearly visible. This would tend to conflict with the previous assumption that the layers are of molecular dimensions but it is thought that the visible film may be the chloride of the metal since it could often be washed off the surface without destroying the hydrophobic properties. On the surface of the unreactive metals such as silver and platinum a film of this type was never formed, although silver

\* The ethyl chlorosilanes were as supplied by Dow Corning Corporation, Midland, Michigan, U.S.A.

occasionally showed a slight staining after extensive treatment. Steel and copper often showed the same characteristics as silver.

The friction tests were all carried out on the Bowden-Leben friction apparatus\*(14) using a load of 4 kg. and a speed of about 0.01 cm./sec. The metals used were copper, mild steel, silver, and platinum coated in the manner described above. In each case the slider was of the same metal as the surface and was not coated with a silicone layer. The friction was measured on the dry silicone films at temperatures ranging from room temperature to 200°–250° C. Similar measurements were made with layers covered with B.P. paraffin oil to test their stability in the presence of mineral oils. An indication of the wear resistant properties of the films was also made by running the loaded slider a number of times over the same track.

The results of these tests are given in Tables 6 and 7.

TABLE 6  
LUBRICATION BY SILICONE FILMS ALONE

Metal	Coefficient of Friction ( $\mu$ )								
	Unlubricated Metal at 20° C.	Monoethyl Silicone		Diethyl Silicone		Mono-isoamyl Silicone		Di-isoamyl Silicone	
		20°C.	c. 200°C.	20°C.	c. 200°C.	20°C.	c. 200°C.	20°C.	c. 200°C.
Copper ..	1.3-1.4	0.45	0.5	0.07	0.10	0.08	0.07	0.10	0.10
Mild steel ..	0.5-0.8	0.2	0.45	0.08	0.10	0.14	0.15	0.14	Very high
Silver ..	1.2-1.4	0.45	0.05	0.12	0.3	0.08	0.08	0.10	0.10
Platinum ..	1.1-1.4	0.5	0.23	0.35	Very high	0.12	Very high	Very high	Very high

\* In this apparatus a small spherical slider (radius 3 mm.) moves over a plane surface at speeds ranging from 0.01 cm./sec. to 1 cm./sec. The spherical upper slider is attached to a circular spring by means of which a normal load up to 8 kg. can be applied between the two surfaces. The upper slider is also mounted on a torsion arm, the deflection of which gives a measure of the frictional force. The lower plane surface can be electrically heated and is fitted with a contact thermocouple. The deflections of the torsion arm and the thermocouple galvanometer are recorded by optical beams on a moving film camera giving a continuous friction-temperature record.



TABLE 7

LUBRICATION BY SILICONE FILMS COVERED WITH PARAFFIN OIL

Metal	Coefficient of Friction ( $\mu$ )								
	Unlubricated Metal at 20°C.	Monoethyl Silicone		Diethyl Silicone		Mono-isoamyl Silicone		Di-isoamyl Silicone	
		20°C.	c. 200°C.	20°C.	c. 200°C.	20°C.	c. 200°C.	20°C.	c. 200°C.
Copper ..	1.3-1.4	0.15	0.4	0.03	0.05	0.14	0.10	0.10	0.10
Mild steel ..	0.5-0.8	0.20	0.10	0.07	0.05	0.13	0.13	0.18	0.18
Silver ..	1.2-1.4	0.25	—	0.07	0.3	0.17	Very high	0.10	0.10
Platinum ..	1.1-1.4	0.25	—	0.35	Very high	0.2	Very high	0.3	Very high

## III. DISCUSSION

An oriented silicone layer appears to be effective as a boundary lubricant. It will be seen from the tables that in most cases it has reduced the coefficient of friction to values characteristic of good boundary lubrication.

With steel and copper at room temperature all the silicones except the monoethyl give good lubrication both alone and when covered with paraffin oil, and, in general, the coefficient of friction is not sensibly changed by heating to 200° C. This is the more remarkable since very few normal boundary lubricants function at all at temperatures much greater than 100° C.

The lubrication of silver has always been a difficult problem, since most of the normal boundary lubricants are ineffective. Yet, except for the monoethyl compound, all the silicones give very low coefficients of friction, and even the monoethyl silicone lubricates well at 200° C. On platinum the silicone films give a relatively high coefficient of friction at room temperature, and in most cases lubrication breaks down on heating. With both silver and platinum the presence of paraffin has little effect.

All previous literature has stated that silicone oils are useless as boundary lubricants because of their great chemical stability, a fact which confirms previous observations that it is necessary to have interaction between the lubricant and metal for boundary lubrication. By coating a metal in the manner described

we obtained an oriented layer which is probably attached to the metal surface of the unreactive metals by the residual field of force of the oxygen atoms, while with the more reactive metals additional anchorage may possibly be supplied by some degree of direct chemical reaction between the metal and the active oxygen. This may explain the observed fact that better lubrication is obtained on copper and steel than on silver and platinum.

Some indication of the strength of the attachment of the silicone layer to the metal surface may perhaps be given by the wear resistance of the film. This was measured by repeated sliding of the upper surface over the same track under a load of 4 kg. until a marked increase in the coefficient of friction indicated that boundary lubrication had broken down. Visual examination of the track showed that such an increase in friction was always accompanied by extensive tearing of the sliding surfaces, typical of unlubricated sliding.

Monoethyl and diethyl silicone films, either alone or in the presence of paraffin, show very little wear resistance on any of the four metals, and the coefficient of friction rises after the first run over the track, although the rate of rise is lower with copper and mild steel than with silver or platinum. However, both the amyl silicones, on copper, silver, and platinum, show a remarkable resistance to wear, for example, with mono- and di-isoamyl silicone films on silver there is no increase in the coefficient of friction when the surfaces are slid over the same track 70 times. In Figures 2 (a) and 2 (b), the coefficient of friction is plotted against the number of runs over the same track, for copper and silver.

In Plate 1, photomicrographs of the tracks produced in sliding give an indication of the extent to which silicone films prevent metallic wear on silver surfaces. The fact that in Plate 1, Figures 3 and 4, the original lapping marks on the metals are practically undisturbed shows that the silicone film has borne the full load of the sliding and has almost completely prevented metal to metal contact.

The addition of paraffin to the silicone film appears to destroy the wear resistance on silver but has no effect on the stability of the film on copper and steel. There is, however, the contradictory effect that on platinum the wear resistance of the silicone film is markedly increased by paraffin.

The stability in the presence of a hydrocarbon oil is obviously important in any practical application of silicone films as boundary lubricants. It must be realized, however, that silicone films prepared in this way differ from the normal boundary lubricant films in that they cannot be repaired after wear by adsorption of further silicone molecules from the bulk of the lubricant. No method has as yet been devised for the production of silicone films by direct adsorption. In view of their excellent lubricating properties and thermal stability, however, an attempt to carry out such a process should give interesting results.

Taking the results in general, it appears that the greater the size and number of the organic radicles on the silane molecule, the better are the lubrication and wear properties. In view of this fact, cetyl chlorsilanes are now being examined but as yet it has not been found possible to produce satisfactory cetyl silicone films.

## IV. CONCLUSION

Molecular silicone films produced by the condensation of chlorsilanes on metal surfaces have been found to have better lubricating properties in some ways than the films formed by the usual organic boundary lubricants. The latter films consist of oriented long chain molecules held to the metal surface by a

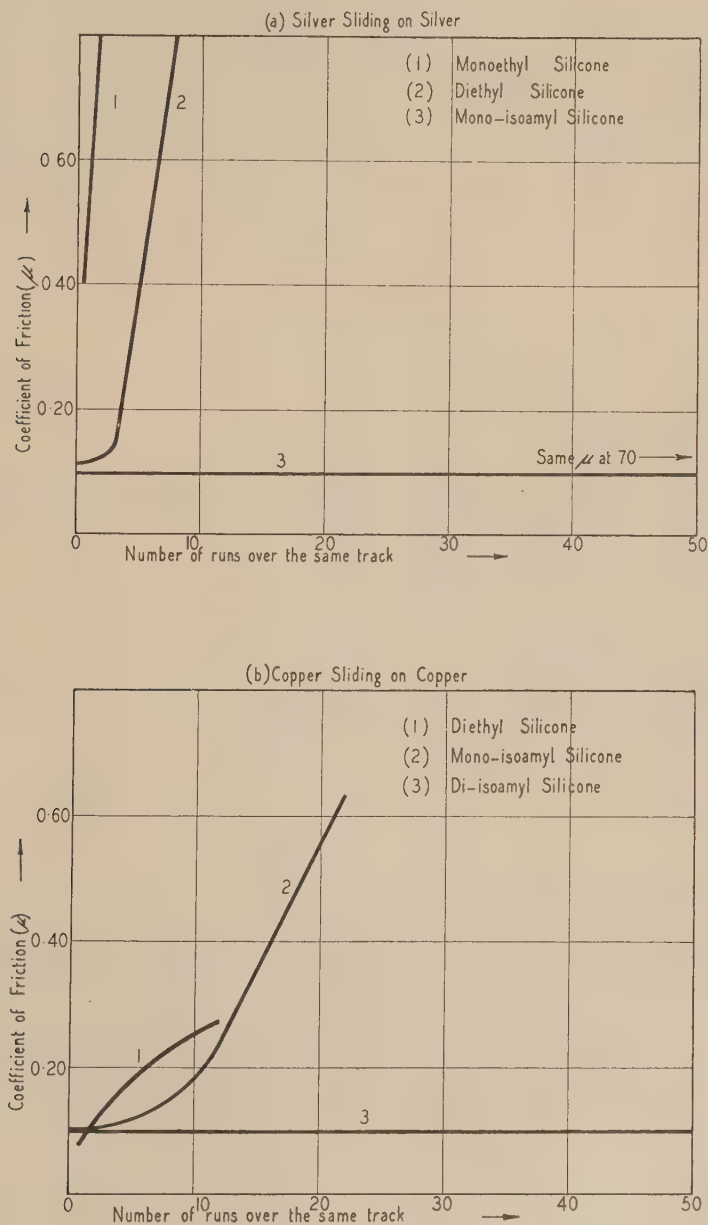


Fig. 2.—Effect of organic radicle on the wearing properties of silicone films.

polar group, with only secondary lateral cohesive forces. Silicone films consist of oriented organic groups attached to each other by the silicon oxygen network or chain, and to the surface by the polarity of the Si-O linkages.

The cross linking in the silicone films is probably responsible for their thermal stability and for the tenacity with which they cling to metal surfaces. On reactive metals, such as copper and steel, it is possible that the silicone would be held at least partly by a direct combination with the atoms of the metal to form a -Si-O-Cu- type of linkage.

## V. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Section of Tribophysics, C.S.I.R.

Our thanks are due to Mr. R. G. Sherwood, Section of Tribophysics, C.S.I.R., who made many of the frictional measurements.

## VI. REFERENCES

- (1) ROCHOW, E. G.—“An Introduction to the Chemistry of the Silicones”, p. 84. (John Wiley & Sons, Inc., 1946.)
- (2) KIPPING, F. S.—*J. Chem. Soc.* **91**: 214 (1907).
- (3) KIPPING, F. S.—*Ibid.* **101**: 2,113 (1912).
- (4) KIPPING, F. S., and ROBISON, R.—*Ibid.* **105**: 484 (1914).
- (5) MEADS, J. A., and KIPPING, F. S.—*Ibid.* **105**: 679 (1914).
- (6) ROCHOW, E. G., and GILLIAM, W. F.—*J. Amer. Chem. Soc.* **63**: 798 (1941).
- (7) HYDE, J. F., and DE LONG, R. C.—*Ibid.* **63**: 1194 (1941).
- (8) ROCHOW, E. G.—*Ibid.* **67**: 963 (1945).
- (9) CHERONIS, N. D., and LEVIN, N.—*J. Chem. Educ.* **22**: 85 (1945).
- (10) KIPPING, F. S.—*J. Chem. Soc.* **123**: 2,598 (1923).
- (11) GILLIAM, W. F., LIEBHAFSKY, H. A., and WINSLOW, A. F.—*J. Amer. Chem. Soc.* **63**: 801 (1941).
- (12) VOGEL, A. I.—“Textbook of Quantitative Inorganic Analysis”, p. 589. (Longmans, Green & Co., Inc., 1941.)
- (13) OLDHAM, J. W. H., and UBBELOHDE, A. R.—*J. Chem. Soc.* **1938**: 201 (1938).
- (14) BOWDEN, F. P., and LEBEN, L.—*Proc. Roy. Soc. A*, **169**: 371 (1939).
- (15) GILMAN, H.—“Organic Syntheses.” Collective Vol. 1, 2nd Ed., p. 25 (1944).

## EXPLANATION OF PLATE 1

Wear of silver surfaces lubricated with silicone films.  $\times 36$ .

- Fig. 1.—No lubricant; 1 run at 20° C.  $\mu=1.2-1.4$ .  
 Fig. 2.—Monoethyl silicone; 1 run at 20° C.  $\mu=0.45$ .  
 Fig. 3.—Diethyl silicone; 1 run at 20° C.  $\mu=0.12$ .  
 Fig. 4.—Mono-isoamyl silicone; 70 runs at 20° C.  $\mu=0.10$ .  
 Fig. 5.—Mono-isoamyl silicone plus paraffin oil; 10 runs at 20° C.  $\mu=0.40$ .  
 Fig. 6.—Di-isoamyl silicone; 70 runs at 20° C.  $\mu=0.04$ .

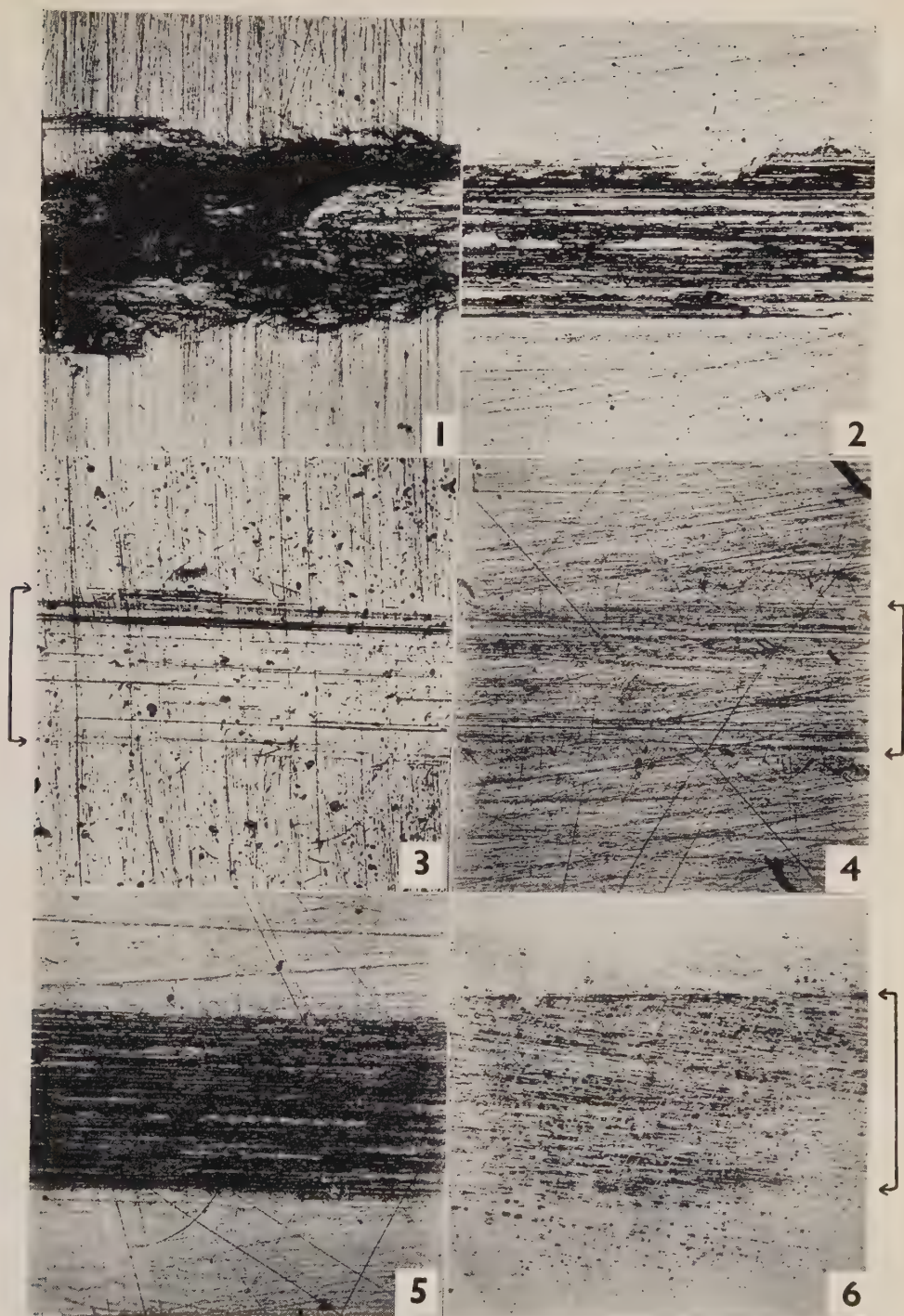
## APPENDIX I

### *Details of Grignard Preparations*

#### (i) *Mono-isoamyl trichlorsilane*

Isoamyl bromide was made by the method given in “Organic Syntheses”(15). One mole isoamyl magnesium bromide in 450 ml. absolute diethyl ether was added







slowly, with stirring, to 113 ml. (1 mole) silicon tetrachloride in 200 ml. absolute ether at room temperature. The stirring was continued for 5 hours after the addition of the Grignard. The mixture was then allowed to stand overnight. After filtering off the magnesium chloride, the ether and unchanged silicon tetrachloride were slowly distilled on a water-bath. Fractional distillation of the residue gave a 50 per cent. yield of fuming colourless liquid boiling between  $158^{\circ}$  and  $162^{\circ}$  C.

When the ratio of  $\text{SiCl}_4$  : Grignard was increased to 1.5 : 1 the yield was 63 per cent.

(ii) *Di-isoamyl dichlorsilane*

(a) One mole isoamyl magnesium bromide in 450 ml. absolute ether was added slowly with stirring to a refluxing solution of 56 ml. ( $\frac{1}{2}$  mole) silicon tetrachloride in 150 ml. absolute ether. The refluxing and stirring were continued for 2 days. The ether was distilled off on a water-bath. Fractionation of the residue yielded 12.5 per cent. mono-isoamyl silane ( $158^{\circ}$ - $162^{\circ}$  C.) and 14 per cent. di-isoamyl silane ( $144^{\circ}$ - $148^{\circ}$  C./56 mm.) plus a higher boiling residue.

(b) One-half mole isoamyl magnesium bromide in 300 ml. absolute ether was added slowly with stirring to a refluxing solution of  $\frac{1}{2}$  mole mono-isoamyl trichlorsilane in 200 ml. ether. The stirring and refluxing were continued for 2 days and then the ether slowly distilled. The yield was not recorded in this case, but an appreciable amount of di-isoamyl dichlorsilane was formed.

(iii) *Tri-isoamyl monochlorsilane*

One-third mole silicon tetrachloride in 150 ml. absolute ether was added slowly with stirring to a refluxing solution of 1 mole isoamyl magnesium bromide in 450 ml. ether. The stirring and refluxing were continued for 5 days and then the mixture allowed to stand for a week. Fractionation yielded 5, 17, and 11 per cent. of mono-, di-, and tri-isoamyl silanes respectively.

---

# "ONE-" AND "TWO-DIMENSIONAL" PARTITION CHROMATOGRAPHIC SEPARATIONS OF ORGANIC ACIDS ON AN INERT SHEET SUPPORT

By J. W. H. LUGG\* and B. T. OVERELL†

(Plate 1)

[Accepted for Publication January 13, 1948]

## Summary

Procedures have been described for "one-" and "two-dimensional" partition-chromatographic separations, on an inert sheet (paper) support, of mixtures of relatively non-volatile organic acids.

Ionization, and adsorption of the acids by the paper, have been suppressed by swamping both the stationary and mobile solvent phases with a volatile acid (acetic or formic). Formic acid is superior in that, at high concentration, it permits well-defined bands of the test acids to travel down the sheets and confers upon the bands highly characteristic  $R_F$  values.

$R_F$  values for some acids (malic, citric, tartaric, succinic, fumaric, malonic, pyruvic, lactic, ketosuccinic (oxalacetic),  $\alpha$ -ketoglutaric, monomethyl succinic, glutaric, and adipic) at 20° C., when Whatman No. 1 filter paper and various pairs of solvent phases are used, have been provided.

Some of the acids present in carrot tissue have been separated in "one-dimensional" work and identified by their  $R_F$  values.

## I. INTRODUCTION

The principles of the so-called "one-dimensional" and "two-dimensional" partition chromatographic separations on inert supports have been discussed in some detail by Consden, Gordon, and Martin(1). For a discussion of the application of these principles to the separation of organic acids, it will be necessary very briefly to describe a highly idealized case.

Of a pair of solvents, between which the various substances to be separated must have distinctly different partition coefficients, one is required to remain stationary in a uniform sheet of suitable, inert material, while the other solvent moves progressively through the material in a predetermined direction. A small quantity of a solution of the substances to be separated is applied to the sheet and is allowed to spread out into a reasonably well-defined spot. The sheet is saturated with the vapour of the stationary phase (i.e. the solvent whose essential properties are that its vapours will provide a thin, virtually immobile layer of liquid on the individual fibres of the sheet without "clogging"). In advance,

\* Biochemistry Department, University of Melbourne.

† Botany Department, University of Melbourne.



this phase has been brought into equilibrium under the prevailing conditions of temperature, etc., with the other phase. The other phase (mobile phase) is then caused to advance as a front along the sheet, and very slowly in comparison with the time required for the continuous adjustments in the equilibria of the partitions of the solutes. As the mobile phase advances past the original spot it abstracts therefrom the individual solutes, and, according to the values of the individual partition coefficients, progressively transports the solutes along the sheet at different rates. Provided that the phase compositions are not appreciably altered by the solutes, that the ratio of the quantity of stationary phase to mobile phase remains the same everywhere through the part of the sheet traversed by the mobile phase, and that the solutes do not undergo marked lateral diffusion in either phase in the course of this process, and provided further that the individual partition coefficients are virtually independent of the concentrations of the solutes, the solutes will be found to occupy separate bands on the sheet. The bands will be of slightly greater lateral dimensions than the original (approximately circular) spot, and will be expanded in the direction of movement if the original spot was very small, or compressed if the spot was relatively large. The excursions of the centres of concentration in the bands from that in the original spot, relative to that of the front of the mobile phase, will be characteristic quantities ( $R_F$  values) independent of the concentrations of the solutes in the original solution.

The relationships have been established by Consden, Gordon, and Martin(1) and may be expressed :

$R_F$  = excursion of band / excursion of mobile phase =  $A_M / (A_M + \alpha A_S)$  where  $\alpha$  is the partition coefficient (stationary phase conc./mobile phase conc.) and  $A_M$  and  $A_S$  are the fractions of a cross-sectional area occupied by the mobile and stationary phases respectively.

The above description of almost idealized events does not apply strictly to actual experimental work. Inert supporting sheets of the requisite uniformity of texture have not yet been employed, and the  $A_S : A_M$  ratio has undoubtedly varied both for this reason and in consequence of the manner in which the mobile phase has been made to advance. These deficiencies, however, have not been serious. Again, if the "stationary" phase were to move a little, then the apparent  $R_F$  values would bear an altered relationship to the partition coefficients, but the relationship could be restored by allowing for the slight movement of this phase. When filter paper is used as the inert support and water comprises the bulk of the stationary phase, the evidence is that the stationary phase is indeed virtually static. But there are other potential difficulties.

Thus, if there is adsorption of the solutes by the material of the "inert" support, then, from the general characteristics of adsorption isotherms of the Freundlich and Langmuir type, the apparent partition coefficients would be expected to increase in favour of the stationary phase accompanying decreases in the concentrations of the solutes. The solutes would be transported along the sheet although they would not be found in well-defined bands but would occupy comet-like areas of the sheet with the "heads" more or less sharply defined and the "tails" more or less long and diffuse, the comets pointing in the

direction of advance of the mobile phase. An entirely similar finding would be expected if the solutes were to suffer a change in partition coefficients in favour of the stationary phase, with decreases in the concentrations. If the partition coefficients were to change in favour of the mobile phase with decrease in concentration, then, in extreme cases, comets would be expected to form and to point back against the direction of advance of the mobile phase.

None of these potential difficulties appears to have been encountered in serious form in the separation of amino-acids(1) or of reducing sugars (Partridge,2). In the work of these authors, filter paper sheets formed the inert support, and water was the chief constituent of the stationary phase. However, Consden, Gordon, and Martin(1) point out that high concentrations of amino-acids can seriously affect the phase compositions. They were able to relate the  $R_F$  values of the amino-acids to their partition coefficients in a manner consistent with the hypothesis that adsorption by the filter paper had been negligible. In any case, the moderately sharp boundaries and the shapes of the individual amino-acid bands suggest either that both adsorption and change in partition coefficients with concentration must have been small factors, or that if they were serious then they must have been mutually self-correcting. But again, there is no *a priori* reason for expecting a marked change in partition coefficients of either sugars or amino-acids with change in concentration over a reasonably wide range of concentrations. The sugars are not appreciably ionized,\* and although on the basis of "zwitter-ion" theory the amino-acids are held to be virtually completely ionized at about neutrality (as regards the  $\alpha$ -NH<sub>2</sub> or =NH and  $\alpha$ -COOH groups), each will ordinarily exist almost exclusively as one species from quite high concentrations to very low concentrations, and over a fairly wide range of pH values. Some difficulty might be anticipated with the dicarboxylic monoamino-acids and with the basic amino-acids, at low concentrations. Consden, Gordon, and Martin(1) did in fact find that the  $R_F$  values for such amino-acids were substantially modified by mild alkalinity or acidity of the phases, but the "comet" effect was not observed. Interesting examples of the inverted comet effect are observed with thyroxine, diiodotyrosine, and thyroxine-o-methyl ether when the phases (from water and n-butanol) are made very strongly alkaline with ammonia (Hird and Trikojus, personal communication 1947).

When an attempt is made to separate organic acids (such as malic, tartaric, citric, succinic, and fumaric) by these means, using a stationary phase composed largely of water in a paper-sheet support, and a neutral mobile phase such as water-saturated n-butanol, a series of overlapping comets, pointing in the direction of movement, is obtained. The individual acids themselves give comets which are more or less characteristic of the acids concerned. Such work as was performed along these lines suggested that the tail length increased with increase in the value of the first ionization constant of the acid in water, and

\* It is to be noted that Partridge(2) obtained markedly elongated spots for a uronic acid when 5-collidine was the chief constituent of the mobile phase.

that the general excursion of the acid decreased with decrease in concentration in the original solution.

These phenomena were briefly discussed by us in a recent publication (Lugg and Overell,<sup>3</sup>). They are to be expected of partially ionized solutes because the degree of ionization will, in general, be higher in the aqueous, stationary phase, and correspondingly, although the partition coefficient of an undissociated species may not change appreciably with concentration, the "gross" partition coefficient (i.e. for undissociated plus dissociated molecules) must be expected to change markedly in favour of the aqueous phase as concentration is decreased. However, the phenomena are also consistent with a hypothesis that there may have been adsorption of the acids by the "inert" support.

There are at least two avenues for seeking a solution of the ionization problem. One is in the direction of obtaining two almost non-ionizing and partially miscible solvents, one of which would remain stationary in a suitable inert support. The other is in the direction of suppressing the ionization of the acids in what would otherwise be ionizing solvents. We have considered both these avenues, but have explored only the second and with sufficient success to warrant our abandonment of the first, at least for the time being.

From measurements of pH values of the organic acid solutions in the absence and in the presence of filter paper, we have been obliged to conclude that there is appreciable adsorption. One may expect that suppression of ionization of the acid would increase its adsorption, but, if the agent used for suppression is present in high concentration it might be adsorbed so strongly itself that it displaces the acid from the filter paper surface. We are of the opinion, therefore, that the procedure described below (designed to suppress the ionization) also effectively suppresses the adsorption.

## II. EXPERIMENTAL

Although we have done some work with paper strips in a chromatography tank suitable only for "one-dimensional" separations, we have usually employed a tank suitable for "two-dimensional" work. It differs slightly from that used by Consden, Gordon, and Martin(1) in having a removable close-fitting lid sealed by water in an exterior annular trough, and equipped with suitably placed viewing holes. The holes are normally kept closed with stoppers. For the purpose of viewing the progress of the mobile phase the stoppers are lifted as glass windows are slid beneath them. The interior of the tank is kept saturated with vapour from the two phases by hanging strips of filter paper wetted with the stationary phase, and by placing separate containers of the two phases at the bottom of the tank. The tank is kept in a well-lagged box and the temperature of the surroundings is thermostatically controlled within  $\pm 1^\circ \text{C}$ . The trough which holds the supply of mobile phase rests on supports placed near the top of the tank. The sheets have one edge immersed in the trough and hang over glass rails vertically in the tank, the mobile phase being induced to flow by capillarity, as in the procedure of Consden, Gordon, and Martin(1).



Our first successful separations of organic acid mixtures were achieved with water and n-butanol as the chief constituents of the two phases and with about 1 mole of acetic acid per litre of each phase. Somewhat better separations were obtained with twice the concentration of acetic acid. In subsequent work we have frequently used formic acid in place of acetic acid. Its ionization constant in water ( $K=2.0 \times 10^{-4}$ ) is considerably greater than that of acetic acid ( $K=1.8 \times 10^{-5}$ ), and at similar concentrations it is more effective in suppressing the comet effect. As main constituents of the mobile phases we have used n-butanol, "amyl" alcohol,\* methyl isobutyl ketone, diisopropyl ketone mesityl oxide, and benzyl alcohol.

In the later work we have standardized the preparation of the phases. Equal volumes of water and of the other solvent are shaken with sufficient formic or acetic acid to give initially a 1.0, 2.0, 4.0, 5.0, or 6.75 M solution of the acid in the aqueous phase. Esterification of alcohols in formic acid systems may be somewhat rapid and pronounced. It is less so in acetic acid systems. We allow the phases comprising such systems to remain in contact with one another (with frequent shaking) for several days, to permit some approach to equilibrium conditions and thereby ensure reasonable stability of the phases during a chromatographic run.

The sheets (we have used No. 1 Whatman papers, 18 in. by 22 in.) are hung in the tank in contact with vapours of the aqueous phase, but for too short a time to saturate them. From 0.003-0.020 ml. (usually 0.010 ml.) of solutions of organic acids to be separated are then applied from a micropipette at marked points near the tops of the sheets, and are allowed to spread into spots of about 0.75-2.1 cm. diameter (depending upon the volume taken) before the tank lid is replaced. After about 30 min., the mobile phase is introduced into the trough via one of the viewing windows, and the tank is left undisturbed until the mobile phase has advanced a sufficient distance (a period of from 6 to 42 hr., depending upon the phases).

The sheets are then removed from the tank and, after marking the positions of the mobile phase fronts, are dried in order to preserve the test acids in the positions they have taken up. In "two-dimensional" separations, the dried sheets are turned through 90° and are then returned to the tank with the appropriate phases for the second run. Following this, they are dried as before. Special treatments could be devised by which the positions of individual test acids might be revealed, but an important aim in these separations is to reveal the positions of all the test acids. This implies the exploitation of a quality they possess in common, namely, their acid functions, and hence, in the drying of the sheets the "swamping" acid added to the phases must also be removed. It is clear, then, that the swamping acid must be of fairly high volatility, and that the test acids to be located must be of sufficiently low volatility to persist during their stay in the tank and during the drying of the sheets.

\* A mixture of about 90 per cent. of 2-methylbutan-4-ol and about 10 per cent. of 2-methylbutan-1-ol.



Drying may be effected by exposing the sheets to fairly dry room air for several days, or by drawing a current of warm air over them in a drying cabinet (not above 60° C.) for an hour or so. However, traces of swamping acid may survive quite lengthy exposure to the heated air. We prefer in all instances to expose the sheets to warm air in the cabinet (at a temperature of about 40° C.) for an hour or so and then to complete the "drying" process by exposing them to room air of moderate humidity for about one or two days. The volatile acids remaining after a preliminary drying in heated air seem to be more easily displaced from the cellulose by the moisture in the air than by prolongation of the exposure to heated air of correspondingly lowered humidity.

After drying, we spray the sheets with a solution of a suitable pH indicator, dry them at low temperature in the cabinet (usually) and then leave them exposed to the room air. We have used various indicators, but for general use bromocresol green and bromophenol blue are probably the most satisfactory. We employ them at concentrations of 0.04 g. of the acid dyestuff per 95 ml. ethanol plus 5 ml. of water, and convert them almost entirely into their sodium salts with alkali (pH 5.5, tint blue, for bromocresol green; pH 4.8, tint reddish-violet, for bromophenol blue) before spraying. Ordinarily, spraying should be no heavier than is sufficient to show the bands clearly, as the test acids are partly neutralized by the indicator. In the finished sheets the acids are revealed by yellow bands, on a greenish-blue ground with bromocresol green, or on a bluish-violet ground with bromophenol blue.

The procedures are necessarily limited to acids of sufficiently low volatility to remain on the paper during exposure in the chromatography tank and during the subsequent drying operations.

### III. RESULTS

$R_F$  values at 20° C. of the various organic acids for sundry pairs of phases (using Whatman No. 1 filter paper) have been computed from chromatograms of the single acids at concentrations of 20-80  $\mu$ g. per 0.01 ml. They are shown in Table 3. Information concerning the acidities of the phases is shown in Table 2.

Samples of the phase-pairs, in the proportions in which the phases separated, were kept at 20° C. for some time and their acidities were checked at intervals. The results are shown in Table 1 and afford some information of the rates of esterification under these conditions and, by comparison with the data in Tables 2 and 3, some idea of the stability of the phases during the chromatographic runs. In passing, it should be mentioned that the benzyl alcohol phase is lighter than the aqueous phase (initial concentration of formic acid, 6.75 M) when first prepared. After esterification has proceeded for some time the phases invert, and it is in this condition that they have been used. Furthermore, in the presence of swamping acid, mesityl oxide slowly undergoes condensation reactions with the development of a dark brown coloration which impairs the quality of the chromatogram; accordingly, phases containing it should be used immediately after preparation.

Some results of applying the procedures to a variety of organic acids in mixtures are reproduced in the accompanying photographs of chromatograph sheets (Plate 1). The indicator used was bromocresol green. The explanation of Plate 1, read in conjunction with Table 3, describes the conditions under which the chromatograms were obtained. As will be seen from the legends, the volumes of solution applied as initial spots varied from 0.003 to 0.020 ml.

TABLE 1  
VARIATION AT 20° C. WITH TIME, OF CONCENTRATION (MOLES/L.) OF SWAMPING ACIDS IN PAIRS OF PHASES\*

Phase Pair	Time in Days										
	0	1	2	3	4	5	7	15	28	34	210
n-Butanol (acetic acid) ..	1.21	1.21									0.92
Water .. .. .	1.00	0.99									0.73
n-Butanol (acetic acid) ..	2.21	2.21									1.67
Water .. .. .	2.00	1.99									1.41
n-Butanol (formic acid) ..	0.85	0.76	0.70								
Water .. .. .	1.00	0.88	0.80								
n-Butanol (formic acid) ..	1.60	1.30	1.23					1.12			0.93
Water .. .. .	2.00	1.56	1.39					1.28			1.06
n-Butanol (formic acid) ..	3.34	2.47	2.10	1.96		1.80			1.59		1.43
Water .. .. .	4.00	3.14	2.80	2.62		2.34			1.88		1.92
"Amyl" alcohol† (acetic acid) .. .. .	1.96	1.96									1.47
Water .. .. .	2.00	2.00									1.54
"Amyl" alcohol† (formic acid) .. .. .	1.16			1.02							0.62
Water .. .. .	2.00			1.66							0.82
"Amyl" alcohol† (formic acid) .. .. .	2.66					2.00				1.12	1.12
Water .. .. .	5.00					3.43				2.26	2.26
Benzyl alcohol (formic acid) .. .. .	4.30	3.12	2.66		2.41		2.25			2.36	
Water .. .. .	6.75	6.24	5.96		5.76		5.60			5.30	

\* There appears to be no change in acidity with time when the chief constituent of the mobile phase is methyl isobutyl or diisopropyl ketone. This is also true for mesityl oxide over a limited period.

† See footnote in Experimental.

A few incidental topics (the behaviour of oxalic acid, enolization of keto-succinic acid, the occasional appearance of spurious bands, and the results of some preliminary work on carrot tissue) are considered in the discussion.

TABLE 2

INITIAL CONCENTRATIONS (MOLES/L.) AND APPROX. MEAN CONCENTRATIONS, DURING THE CHROMATOGRAPHY RUNS, OF THE SWAMPING ACIDS IN PAIRS OF PHASES

No. of Phase Pair	Phase Pair and Swamping Acid	Initial Concentration (Moles/l.) of Swamping Acid	Approx. Concentration of Swamping Acid (Moles/l.) during Run
1	n-Butanol (acetic acid) ..	1.21	1.21
	Water .. .. .	1.00	0.99
2	n-Butanol (acetic acid) ..	2.21	2.21
	Water .. .. .	2.00	1.99
3	n-Butanol (formic acid) ..	0.85	0.68
	Water .. .. .	1.00	0.78
4	n-Butanol (formic acid) ..	1.60	1.22
	Water .. .. .	2.00	1.37
5	n-Butanol (formic acid) ..	3.34	1.78
	Water .. .. .	4.00	2.30
6	"Amyl" alcohol* (acetic acid)	1.96	1.96
	Water .. .. .	2.00	2.00
7	"Amyl" alcohol* (formic acid)	1.16	1.00
	Water .. .. .	2.00	1.63
8	"Amyl" alcohol* (formic acid)	2.66	1.98
	Water .. .. .	5.00	3.41
9	Benzyl alcohol (formic acid) ..	4.30	2.20
	Water .. .. .	6.75	5.55
10	Methyl-isobutyl ketone (formic acid) .. .. .	2.44	2.44
	Water .. .. .	5.00	5.00
11	Diisopropyl ketone (formic acid)	1.40	1.40
	Water .. .. .	5.00	5.00
12	Mesityl oxide (formic acid) ..	3.60	3.60
	Water .. .. .	5.00	5.00

\* See footnote in Experimental.

TABLE 3

*R<sub>f</sub>* VALUES FOR ORGANIC ACIDS WITH THE PHASE PAIRS LISTED IN TABLE 2 (AT 20° C. AND USING WHATMAN NO. 1 FILTER PAPER) TOGETHER WITH INFORMATION CONCERNING THE IONIZATION CONSTANTS OF THE ACIDS (pK VALUES)

No. of Phase Pair	Malic	Tartaric	Citric	Succinic	Fumaric	Malonic	Pyruvic	Lactic	Keto-succinic	$\alpha$ -Keto-glutaric	Mono-methyl succinic	Glutaric	Adipic
	pK <sub>1</sub> 3.48* pK <sub>2</sub> 5.11*	pK <sub>1</sub> 3.00* pK <sub>2</sub> 4.39*	pK <sub>1</sub> 3.08* pK <sub>2</sub> 4.39* pK <sub>3</sub> 5.49*	pK <sub>1</sub> 4.18* pK <sub>2</sub> 5.57*	pK <sub>1</sub> 3.03* pK <sub>2</sub> 4.49*	pK <sub>1</sub> 2.80* pK <sub>2</sub> 5.68*	pK 2.22*	pK 3.85*	pK <sub>1</sub> ** pK <sub>2</sub> **	pK <sub>1</sub> ** pK <sub>2</sub> **	pK <sub>1</sub> 4.07* pK <sub>2</sub> 5.78*	pK <sub>1</sub> 4.72* pK <sub>2</sub> 5.57*	pK <sub>1</sub> 4.43* pK <sub>2</sub> 5.62*
1	0.45† 0.48 0.53†	0.20† 0.26 0.28†	0.32† 0.37 0.41†	0.83† 0.78 0.80†	0.85† 0.84								
2	0.50 0.52†	0.35 0.38†	0.44 0.46†	0.71 0.72†	0.79 0.81†								
3	0.46 0.47†	0.24 0.25†	0.40 0.40†	0.70 0.70†	0.84 0.84†								
4	0.46	0.26	0.41	0.70	0.84								
5	0.44† 0.45 0.45†	0.21† 0.22 0.22†	0.37† 0.37 0.39†	0.74† 0.74 0.75†	0.90† 0.89 0.88†	0.63	0.66	0.67	0.64	0.58	0.86	0.84	0.89
6	0.26† 0.26 0.33†	0.09† 0.11 0.14†	0.15† 0.17 0.24†	0.62	0.70† 0.71 0.76†								
7	0.29† 0.31 0.32†	0.10† 0.12 0.13†	0.19† 0.21 0.22†	0.62† 0.64 0.64†	0.74† 0.75 0.76†								



TABLE 3—Continued

No. of Phase Pair	Malic	Tartaric	Citric	Succinic	Fumaric	Malonic	Pyruvic	Lactic	Keto-succinic	$\alpha$ -Keto-glutaric	Mono-methyl succinic	Glutaric	Adipic
	pK <sub>1</sub> 3.48* pK <sub>2</sub> 5.11*	pK <sub>1</sub> 3.00* pK <sub>2</sub> 4.39*	pK <sub>1</sub> 3.08* pK <sub>2</sub> 4.39* pK <sub>3</sub> 5.49*	pK <sub>1</sub> 4.18* pK <sub>2</sub> 5.57*	pK <sub>1</sub> 3.03* pK <sub>2</sub> 4.49*	pK <sub>1</sub> 2.80* pK <sub>2</sub> 5.68*	pK 2.22*	pK 3.85*	pK <sub>1</sub> ** pK <sub>2</sub> **	pK <sub>1</sub> ** pK <sub>2</sub> **	pK <sub>1</sub> 4.07* pK <sub>2</sub> 5.78*	pK <sub>1</sub> 4.72* pK <sub>2</sub> 5.57*	pK <sub>1</sub> 4.43* pK <sub>2</sub> 5.62*
8	0.33† 0.33 0.33†	0.12† 0.13 0.13†	0.22† 0.23 0.23†	0.71† 0.70 0.69†	0.87† 0.84 0.84†	0.54							
9	0.19† 0.19 0.20†	0.06† 0.06 0.06†	0.12† 0.12 0.12†	0.57† 0.56 0.57†	0.71† 0.71 0.69†	0.41	0.49†	0.47†	0.42	0.37			
10	0.18† 0.18 0.17†	0.06† 0.05 0.05†	0.11† 0.10 0.10†	0.58† 0.59 0.58†	0.81† 0.81 0.81†	0.50	0.60	0.52					
11	0.04 0.04†	<0.01	<0.01	0.20	0.52† 0.52 0.52†	0.13			0.15				
12	0.37† 0.37 0.35†	0.17† 0.16 0.15†	0.30† 0.30 0.28†		0.92 0.90†								

\* pK values taken from Clark(4) and Beilstein(5).

\*\* pK values for the keto form of ketosuccinic are not available (enol forms: hydroxynaleic pK<sub>1</sub>=2.60, hydroxyfumaric pK<sub>1</sub>=2.56).† pK values for  $\alpha$ -ketoglutaric are not available.‡ *R<sub>F</sub>* values for 20  $\mu$ g. of the acids in 0.01 ml.‡ *R<sub>F</sub>* values for 80  $\mu$ g. of the acids in 0.01 ml.Other *R<sub>F</sub>* values for 40  $\mu$ g. of the acids in 0.01 ml.

## IV. DISCUSSION

It is clear that the phases must have contained considerable quantities of esters in some instances (notably when *n*-butanol and "amyl" alcohol were used with formic acid). Apart from the effects of the swamping acids upon the ionizations and adsorptions of the test acids (suggested by the progressive elimination of comet effect with increasing acidity of the phases, if our interpretation is sound), there were the effects of the swamping acids and of esters upon the phase compositions, one of which was an appreciable increase in concentration of the major component of the stationary in the mobile phase, and vice versa. Consequently, modification of  $R_F$  values by greater or lesser concentration of swamping acid cannot be interpreted entirely in terms of the effectiveness of suppression of ionization and adsorption of the test acids. That the band excursions were increased by the suppression may reasonably be presumed. Confirmation is afforded by the fact that the band excursion of a test acid for any particular pair of phases was usually increased by the presence of other test acids and by increase in the concentration, and that these effects were most marked when the pH of the aqueous phase was highest (low concentrations of acetic acid), and least marked, and sometimes negligible, when it was lowest (high concentrations of formic acid).

Again, in this connection, it is to be recorded that, in general, the comet effect was most marked with test acids which (judging from their ionization constants) must have been the most extensively ionized. Thus, while the degrees of ionization of all the test acids in dilute solution must have been fairly extensive in the stationary phase in absence of swamping acid (e.g. approximately 0.04 for succinic and 0.2 for malonic at concentrations of 40  $\mu$ g. per 0.01 ml., and increasing greatly as the acid was withdrawn from the spot), the extents of suppression of their ionizations by the swamping acid must have been most marked for the test acids with lowest ionization constants (degrees of ionization in 2.3 M formic acid, e.g. approximately 0.003 for succinic acid and 0.08 for malonic acid, and changing very little as the acid was withdrawn from the spot). At the higher concentrations of formic acid as swamping acid, the comet effect was virtually eliminated (see Plate 1) for all test acids (except oxalic acid, discussed later), and the  $R_F$  values were almost independent of the concentrations of test acids individually (see Table 1) and of the presence of other test acids in mixtures.

Unfortunately, although the ratios of  $R_F$  values changed, the order of separation of the test acids proved to be the same with different pairs of solvent phases. This circumstance seriously limits the effectiveness of "two-dimensional" resolutions of mixtures of such test acids. However, mixtures which contain some acids with aromatic groups might be resolved effectively in "two-dimensional" work by using benzyl alcohol as the main constituent of the mobile phase for one of the chromatographic runs.

Plate 1, Figure 4, showing a "two-dimensional" separation of a test mixture, reveals that the  $R_F$  values for the second chromatographic run have not been modified appreciably by the first run, but depending upon the mobile phases selected and the order in which they are used, the  $R_F$  values for the

second run are sometimes affected. This effect is presumably attributable to a change in the qualities of the paper occasioned by the first run and the subsequent drying of the sheet. For serious "two-dimensional" work it will be necessary, therefore, to determine a new series of  $R_F$  values for paper sheets thus modified.

As will be seen from Plate 1 and from Table 3, certain relationships are discernible between the  $R_F$  values and the characteristics of the test acids. Thus the  $R_F$  values tend to be diminished by (i) increase in the number of carboxyl groups, (ii) increase in the number of non-carboxylic hydroxyl groups, (iii) presence of a non-carboxyl carbonyl group, and (iv) decrease in the hydrocarbon character (i.e. in the number of  $-\text{CH}=\text{}$ ,  $-\text{CH}_2-$ , and  $-\text{CH}_3$  groups). These effects, which were expected, may be of value in indicating the probable character of an unknown acid. In this connection it should be mentioned that the excursion of oxalic acid was always found to be very small, but unfortunately, even with the highest concentrations of formic acid in the phases, its comet effect (in keeping with its strongly acid nature:  $K_1=6.5 \times 10^{-2}$ ,  $K_2=6.1 \times 10^{-5}$ ) was still marked. Approximate  $R_F$  values of oxalic acid with phase-pairs 8 and 11 were 0.10 and 0.06 respectively.

Ketosuccinic (oxalacetic) acid is subject to enolization and may exist as ketosuccinic, hydroxymaleic, or hydroxyfumaric acid, or as mixtures of these. From the manner of its preparation (crystallization from fairly concentrated hydrochloric acid) and of its use in the chromatographic work (the phases were strongly acidified with formic acid), it is presumed that the acid was present as ketosuccinic acid.

The occasional appearance of a weak "ghost" band in the position of the initial spot and/or of one or two weak bands a little distance beneath the spot, remains an unexplained phenomenon. These bands may be due to the formation, under the influence of the test acid(s) in the initial spot, of acid (uronic?) residues from glucose residues of the cellulose fibres, followed by partial liberation of (uronic?) acid and (uronic?) acid complexes. A point, however, which tends to invalidate this attempted explanation, is that the supposed effect upon the cellulose appears to be confined to the test acid(s) while they are in the initial spot.

In the earlier work in which acetic acid was used as swamping acid, we found it possible to estimate roughly the quantities of test acids by excising the areas of paper enclosing the bands, steeping the areas in water, and titrating with alkali. There is evidence that fractions of the test acids are missing from the bands when formic acid is used as swamping acid. This effect, which may be due to partial esterification of the test acids by alcohols in the strongly acid phases and/or to partial esterification of test acid hydroxyl groups by swamping acid, is under investigation.

By "one-dimensional" chromatography of dialysates of formic acid extracts of carrot tissue, we have demonstrated the presence of malic, fumaric, and succinic acids (or at least of their salts) in the tissue. These acids were identified by their  $R_F$  values. The strips were treated also with "ninhydrin" (see Consden, Gordon, and Martin, 1) to reveal the positions of amino-acids or peptides which might otherwise have vitiated the interpretations, but these substances appeared

to be confined to the upper portions of the strips. In more searching investigations of the organic acids in plant and animal tissues it will be advantageous to extract them from the acidified tissues with ether or some other suitable solvent.

## V. ACKNOWLEDGMENTS

We are indebted to Mr. P. Stanley for the sample of ketosuccinic (oxalacetic) acid and to Mr. J. F. Turner, Division of Food Preservation and Transport, C.S.I.R., for the samples of pyruvic and  $\alpha$ -ketoglutaric acids. Thanks are due also to Mr. G. Ogilby for photographing the chromatograms.

## VI. REFERENCES

- (1) CONSDEN, R., GORDON, A. H., and MARTIN, A. J. P.—*Biochem. J.* **38**: 224 (1944).
- (2) PARTRIDGE, S. M.—*Nature* **158**: 270 (1946).
- (3) LUGG, J. W. H., and OVERELL, B. T.—*Ibid.* **160**: 87 (1947).
- (4) CLARK, W. M.—“The Determination of Hydrogen Ions.” (Baillière, Tindall and Cox: London, 1928.)
- (5) BEILSTEIN, F. K.—“Handbuch der organischen Chemie”, 4th Ed. (Deut. Chem. Ges.: Berlin, 1921.)

## EXPLANATION OF PLATE 1

Chromatograms illustrating the application of the methods to the separation of a variety of organic acids.

The band positions of the individual test acids are indicated by the following abbreviations: Ma (malic), T (tartaric), C (citric), Mo (malonic), KS (ketosuccinic=oxalacetic), La (lactic), F (fumaric), S (succinic), KG ( $\alpha$ -ketoglutaric).

The mobile phase fronts are indicated in the “one-dimensional” chromatograms (Figs. 1-3) by dark lines near the bottom edges of the strips. In the “two-dimensional” chromatogram the fronts are indicated in Figure 4.

Fig. 1.—Chief constituent of mobile phase, n-butanol.

- First strip.* No swamping acid, 40  $\mu$ g. of each test acid (T, C, Ma, S, and F) in 0.01 ml.
- Second strip.* Initial conc. of acetic acid in aq. phase, 1.00 M, 40  $\mu$ g. of each test acid in 0.02 ml.
- Third strip.* Initial conc. of formic acid in aq. phase, 1.00 M, 40  $\mu$ g. of each test acid in 0.02 ml.
- Fourth strip.* Initial conc. of formic acid in aq. phase, 2.00 M, 40  $\mu$ g. of each test acid in 0.01 ml.
- Fifth strip.* Initial conc. of formic acid in aq. phase, 4.00 M, 40  $\mu$ g. of each test acid in 0.01 ml.

Fig. 2.—Chief constituent of mobile phase for the first three strips, n-butanol; initial conc. of formic acid in aq. phase, 4.00 M.

- First strip.* 80  $\mu$ g. of each test acid in 0.01 ml.
- Second strip.* 40  $\mu$ g. of each test acid in 0.01 ml.
- Third strip.* 20  $\mu$ g. of each test acid in 0.01 ml.

Chief constituent of mobile phase for the last three strips, “amyl” alcohol; initial conc. of formic acid in aq. phase, 5.00 M.

- Fourth strip.* 80  $\mu$ g. of each test acid in 0.01 ml.
- Fifth strip.* 40  $\mu$ g. of each test acid in 0.01 ml.
- Sixth strip.* 20  $\mu$ g. of each test acid in 0.01 ml.





LUGG and OVERELL.—“ONE-” AND “TWO-DIMENSIONAL” PARTITION CHROMATOGRAPHIC SEPARATIONS OF ORGANIC ACIDS ON AN INERT SHEET SUPPORT



Fig. 3.

- First strip.* Chief constituent of mobile phase, n-butanol; initial conc. of formic acid in aq. phase, 4.00 M; 40  $\mu$ g. of each test acid in 0.01 ml.
- Second strip.* Chief constituent of mobile phase, "amyl" alcohol\*; initial conc. of formic acid in aq. phase, 5.00 M; 40  $\mu$ g. of each test acid in 0.01 ml.
- Third strip.* Chief constituent of mobile phase, benzyl alcohol; initial conc. of formic acid in aq. phase, 6.75 M; 18  $\mu$ g. of each test acid in 0.003 ml.
- Fourth strip.* Chief constituent of mobile phase, n-butanol; initial conc. of formic acid in aq. phase, 4.00 M; 18  $\mu$ g. of each test acid in 0.003 ml.

\* See footnote in Experimental.

Fig. 4.—"Two-dimensional" chromatogram, 40  $\mu$ g. of each test acid in 0.01 ml.

- First run.* Chief constituent of mobile phase, n-butanol; initial conc. of formic acid in aq. phase, 4.00 M.
- Second run.* Chief constituent of mobile phase, benzyl alcohol; initial conc. of formic acid in aq. phase, 6.75 M.

# STUDIES ON THE LIGNIN OF *EUCALYPTUS REGNANS*

## I. THIOLIGNIN

By F. N. LAHEY\* and J. W. T. MEREWETHER†

[Accepted for Publication November 26, 1947]

### Summary

By pulping *Eucalyptus regnans* by the kraft process, acidifying the black liquor, and fractionating the product, a thiolignin has been prepared in a pure state.

Thiolignin reacts with diazomethane to yield a trimethyl derivative, and with dimethyl sulphate to yield a hexamethyl derivative.

Thiolignin, trimethylthiolignin, and hexamethylthiolignin react with acetic anhydride in pyridine to yield respectively hepta-, tetra-, and mono-acetyl derivatives.

The data obtained are consistent with the formula



## I. INTRODUCTION

The experiments reported in the present paper deal with thiolignin, an alkali-soluble derivative of lignin containing sulphur, produced during the pulping of wood by the sulphate process and isolated by acidifying the waste liquors from this process. In normal commercial practice it is burned during the process of recovering the alkali, and it is perhaps for this reason that comparatively little attention has been paid to the problem of determining the chemical properties and constitution of thiolignin, or to the mechanism of the reactions taking place between lignin and the chemicals used in this process.

In the manufacture of wood pulp by the soda process, wood chips are digested at an elevated temperature with a solution of sodium hydroxide, and most of the lignin in the wood is converted into alkali-soluble derivatives known as alkali-lignins. In the kraft or sulphate process 20–30 per cent. of the total alkali is replaced by sodium sulphide and the lignin is converted into a mixture of alkali-lignins and thiolignin. Klason and Segerfeld(1) showed that acidification of the black liquor obtained in the kraft pulping of spruce yielded a product containing combined sulphur, and it was demonstrated by Anderzen and Holmberg(2), and later by Cinves(3) and Hägglund(4), that the product was a mixture. Hägglund and Lumme(5), on the basis of their observation that the consumption of sodium sulphide was related to the ratio of sulphide to hydroxide, inferred that the amount of sulphur taken up by the lignin was not fixed but

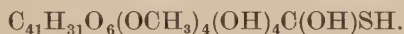
\* Chemistry Department, University of Melbourne.

† Research Laboratories, Australian Paper Manufacturers Limited, Melbourne.



depended on the conditions used in pulping. This was subsequently shown to be correct by Hanson(6). That other changes may take place was indicated by Kimble(7), who found that partial demethylation may take place during kraft pulping. Interesting data on the ease of solution of both lignin and thioglignin in sodium hydrosulphide and sodium hydroxide have been provided by Kullgren(8), Hägglund(4), and Brauns(9).

The most detailed investigation on this subject is that of Ahlm(10), who prepared a series of derivatives from a carefully purified spruce thioglignin, and on the basis of his analytical data suggested the formula



It is well known that hardwood lignin differs from softwood lignin, and the object of the work described in the present paper was to obtain information as to the nature of the lignin isolated from the black liquor obtained when *Eucalyptus regnans* wood is pulped under conditions similar to those used in commercial practice.

## II. THIOLIGNIN FROM *E. regnans*

Preliminary experiments (Table 1) on the acidification of black liquor under varying conditions of temperature and pH confirmed the claim of Cooke(11) that the highest yield of crude lignin is obtained by acidifying at 50° C. and a

TABLE I  
YIELD OF CRUDE LIGNIN (G./LITRE) OBTAINED BY ACIDIFYING  
BLACK LIQUOR WITH 50 PER CENT.  $\text{H}_2\text{SO}_4$

pH	Temperature		
	20° C.	50° C.	80° C.
9	g./l. 37	g./l. 29	g./l. 28
5.5	50	63	39
2	68	80	43

low pH, and these conditions were accordingly used throughout. Free sulphur was removed from this crude lignin by Soxhlet extraction with carbon disulphide, after which it was dissolved in anhydrous dioxan and filtered free from carbohydrate impurities.

Addition of this dioxan solution to an excess of anhydrous ether precipitated the thioglignin, and this operation was repeated until the methoxyl content reached a value which was unchanged by further treatment. As a further check

on its purity a sample was "recrystallized" from cyclohexanol, yielding a product with unchanged methoxyl, sulphur, and equivalent weight. The fraction soluble in dioxan-ether contained no sulphur and presumably consisted of alkali-lignins.

Potentiometric titration of the thiolignin (Fig. 1) showed two points of inflexion on the curve and indicated an equivalent weight of 665. Analysis gave the figures: C, 61.5%; H, 5.7%; S, 2.4%. An approximate minimum molecular formula to fit these data is  $C_{68}H_{76}O_{25}S$  (mol. wt. 1324). That the molecule is of this order of magnitude was shown by a molecular weight determination (Rast) of its hexamethyl derivative whereby a figure of 1420 was obtained, a value in close agreement with the calculated figure of 1408.

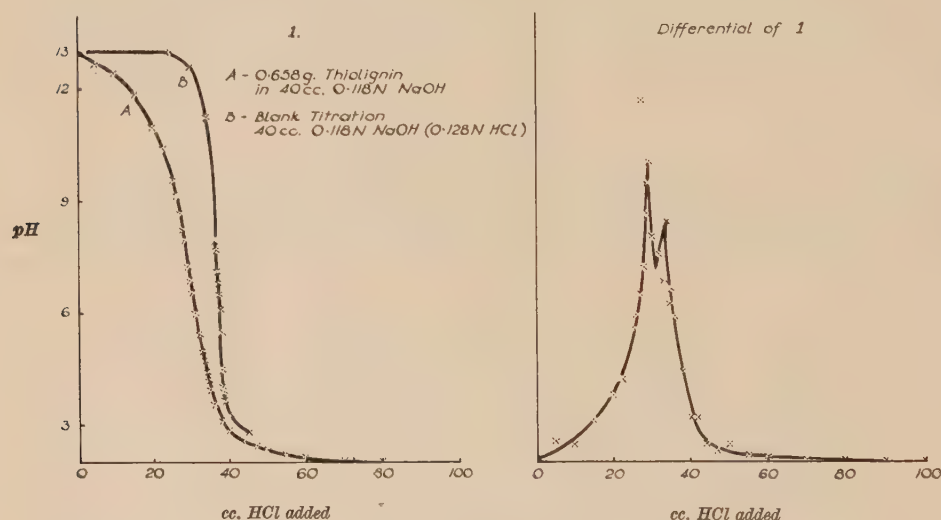


Fig. 1.—Potentiometric titration of thiolignin.

The methoxyl content of thiolignin is 21.1 per cent. The number of methoxyl groups in the molecule can be calculated from the equation

$$\frac{a \times 31.03}{1324} = \frac{21.1}{100}$$

whence  $a=9$ .

When thiolignin is treated with acetic anhydride in pyridine there is formed an acetyl derivative containing 18.6 per cent. acetyl. The total number of hydroxyl groups will therefore be given by the equation

$$\frac{b \times 43.04}{1324 + 42b} = \frac{18.6}{100}$$

whence  $b=7$ .

The thiolignin molecule thus contains nine methoxyl and seven hydroxyl groups. In order to obtain more precise information as to the latter, thiolignin

was treated with diazomethane, yielding a product containing 27.1 per cent. methoxyl, and from the equation

$$\frac{(9+c) \times 31.03}{1324 + 14c} = \frac{27.1}{100}$$

the number of acidic hydroxyl groups is found to be three. On acetylation this yields a tetra-acetate.

These three acidic hydroxyls could be either carboxylic, phenolic, or enolic, or a corresponding thio-analogue. That there is no carboxyl group is shown by the fact that the three react equally with acetic anhydride and with diazomethane.

The presence of an enol was shown by a positive reaction to the Br<sub>2</sub>-HI test for enols, and in a subsequent paper it will be shown that there is only one enol present.

In view of the presence of sulphur it appeared probable that one of the groups to react with diazomethane would be a mercaptan. This was supported firstly by the fact that thioglignin is not only soluble in alkali but also in dilute aqueous sodium bicarbonate. Secondly, when an alcoholic solution of thioglignin was treated with alcoholic lead acetate all the thioglignin was precipitated. Finally, a positive reaction was obtained to Feigl's test(12), the catalysis of the reaction  $2\text{NaN}_3 + \text{I}_2 = 2\text{NaI} + 3\text{N}_2$ , a test which is stated to be specific amongst organic compounds for mercaptans and thioketones.

Methylation with dimethyl sulphate and 30 per cent. sodium hydroxide, the method used by most workers in lignin chemistry to obtain complete methylation, did not achieve this result. By simultaneous deacetylation and methylation either of heptacetylthioglignin or of tetracetyltrimethylthioglignin it was found possible to methylate only three of the four alcoholic hydroxyls. Prolonged treatment either had no further effect, or resulted in demethylation, the methoxyl content dropping from 33.0 to 26.5, 26.7, and 27.8 per cent. in three different attempts. By methylation with dimethyl sulphate in acetone in the presence of anhydrous potassium carbonate similar results were obtained, attempts to methylate the last hydroxyl group resulting only in demethylation. That this hydroxyl group was unchanged by the treatment was shown by the fact that hexamethylthioglignin could be acetylated with acetic anhydride in pyridine to yield a mono-acetate.

In a subsequent paper it will be shown that this hydroxyl group is a tertiary alcohol. Freudenberg and Sohns(13), in one of their investigations concluded that the lignin they were working on contained a tertiary hydroxyl, and found that C-methyl groups were present to the same extent, inferring from that the presence of the  $:\text{C}(\text{CH}_3)\text{OH}$  grouping. Thioglignin was found to contain 1.1 per cent. C-CH<sub>3</sub>, equivalent to one C-methyl group.

On treating trimethylthioglignin with triphenylchloromethane in pyridine no reaction took place, unchanged trimethylthioglignin being recovered, indicating the absence of primary alcoholic hydroxyls. It is unlikely that thioglignin contains a double bond, as diazomethane usually reacts with such to yield pyrazolines, and no nitrogen was found in trimethylthioglignin.

Summarizing these data, thiolignin prepared from *Eucalyptus regnans* in the manner described has been found to contain nine methoxyls, one thiol, one enolic hydroxyl, one phenolic hydroxyl (by difference), three secondary alcoholic hydroxyls, one tertiary alcoholic hydroxyl, and one methyl attached to carbon, and the evidence obtained agrees with the formula



### III. EXPERIMENTAL

#### (a) *Crude lignin*

Six logs of green 21-year-old *Eucalyptus regnans* from Powelltown, Victoria, were converted into chips; these chips were then screened to remove oversize and fines, sorted to remove gum veins, etc., and air-dried. In a steam-jacketed rotary iron digester were mixed chips (4,000 g. dry weight), kraft cooking liquor of 22 per cent. sulphidity (the ratio of sodium sulphide to sodium hydroxide plus sulphide, both expressed as NaOH) equivalent to 800 g. active NaOH, and water to make the total volume (including cooking liquor and moisture in wood) to 16 litres. The charge was taken up to boiling point in 45 minutes, when all the air was blown out, to 166° C. in another 45 minutes, and maintained at that temperature for a further 2½ hours. On cooling, the spent liquor was filtered on a Buchner and the residual pulp washed three times with water.

The combined filtrate and washings were filtered once more, warmed to 50°–55° C. and 50 per cent. H<sub>2</sub>SO<sub>4</sub> added slowly with stirring until the mixture was acid to Congo. During the addition vigorous foaming occurred as a result of the copious evolution of SO<sub>2</sub> and other gases, whilst the colour changed from black to light brown as the crude lignin was precipitated. This was allowed to stand for one hour, after which the supernatant liquor (a fine suspension) was decanted off, fresh water added, and the lignin filtered under suction and washed with water. The crude lignin was then dried for three days on the top of an oven. Yield 478 grams (12 per cent. on the original wood).

The crude lignin is a reddish-brown powder with no sharp melting point; it commences to soften at 122° C. and finally liquefies at 176° C.

#### (b) *Thiolignin*

The crude lignin was extracted for 48 hours in a Soxhlet extractor with carbon disulphide to remove free sulphur, and the extracted lignin (100 g.) dissolved in anhydrous dioxan (1,000 cc.), warmed, and filtered while hot. On cooling, the filtered dioxan solution was added dropwise to anhydrous ether (10 litres) with vigorous stirring. The precipitated thiolignin was washed repeatedly with anhydrous ether until the washings were colourless, and air-dried. Yield 67 grams.

This purification was then repeated four times, each time dissolving the thiolignin in dioxan to give a 10 per cent. w/v solution and using ten volumes of anhydrous ether. After the final precipitation, the methoxyl content was unchanged. Final yield 53 grams,



Found: C, 61.5; H, 5.7; S, 2.3;  $\text{OCH}_3$ , 21.1;  $\text{C}(\text{CH}_3)$ , 1.1%; eq. wt. 662.

Calculated for  $\text{C}_{68}\text{H}_{76}\text{O}_{25}\text{S}$  (1324): C, 61.6; H, 5.7; S, 2.4;  $\text{OCH}_3$ , 21.1;  $\text{C}(\text{CH}_3)$ , 1.1%; eq. wt. 662.

A sample was repurified from boiling cyclohexanol.

Found: S, 2.4;  $\text{OCH}_3$ , 21.1%; eq. wt. 663.

Thiolignin is a fine light-brown powder which has no sharp melting point, but which commences to soften at  $230^\circ\text{C}$ . It is soluble in dioxan, pyridine, chloroform, acetone, glacial acetic acid, boiling cyclohexanol, and aqueous sodium bicarbonate. It is slightly soluble in ethyl alcohol and cold cyclohexanol, and insoluble in water, benzene, ether, petroleum ether, and carbon disulphide.

### (c) *Trimethylthiolignin*

Anhydrous dioxan (90 cc.) was added to 50 per cent. aqueous potassium hydroxide (30 cc.) and the mixture cooled in ice until the dioxan just started to freeze. Nitrosomethylurea (10 g.) was then added in small portions, shaking well between additions, and maintaining the temperature just above the freezing point of the dioxan. The yellow dioxan solution of diazomethane was decanted off and the reaction mixture shaken with four successive portions of 30 cc. dioxan. The combined solutions were dried for three hours over KOH pellets and then added to a solution of thiolignin (20 g.) in anhydrous dioxan (200 cc.). The mixture was allowed to stand at room temperature for 48 hours.

A small portion was isolated for analysis by filtering, adding dropwise to ten volumes of ether, and purifying twice from dioxan-ether. The remainder was concentrated to about 150 cc. and remethylated under the same conditions. The process was repeated a third time, but this was found to be unnecessary, the methoxyl content remaining unaltered. The product was isolated by filtering off the diazomethane polymer, concentrating to 200 cc., and adding dropwise to ten volumes of anhydrous ether. Yield 19.1 grams;  $\text{OCH}_3$ , 27.6%; S, 1.8%. It was then purified five times from dioxan-ether. Yield 15.5 grams.

Found: C, 62.4; H, 6.1; S, 2.2;  $\text{OCH}_3$ , 27.1%.

Calculated for  $\text{C}_{71}\text{H}_{82}\text{O}_{25}\text{S}$  (1366): C, 62.3; H, 6.0; S, 2.3;  $\text{OCH}_3$ , 27.2%.

Trimethylthiolignin is a light-buff powder, has no sharp melting point, but commences to soften and darken in colour at  $170^\circ\text{C}$ . It is soluble in dioxan, pyridine, chloroform, and acetone, and insoluble in water, aqueous alkali, ethyl alcohol, ether, benzene, and petroleum ether.

### (d) *Heptacetylthiolignin*

Thiolignin (20 g.) was dissolved in anhydrous pyridine (100 cc.), and acetic anhydride (60 cc.) added. After standing at room temperature for 48 hours the solution was added dropwise to anhydrous ether (1,500 cc.) with vigorous stirring. The crude product was precipitated as a fine powder which was filtered off and washed with anhydrous ether. Yield 21.1 grams. This was purified by dissolving in anhydrous dioxan (100 cc.) and adding dropwise to anhydrous ether

(1,000 cc.), filtering, and washing with anhydrous ether. After five purifications the yield was 17.7 grams.

Found: C, 61.0; H, 5.6; S, 2.2;  $\text{OCH}_3$ , 18.4;  $\text{COCH}_3$ , 18.6%.

Calculated for  $\text{C}_{82}\text{H}_{90}\text{O}_{32}\text{S}$  (1618): C, 60.8; H, 5.6; S, 2.0;  $\text{OCH}_3$ , 17.9;  $\text{COCH}_3$ , 18.6%.

Heptacetylthiolignin is a light-buff powder soluble in the same solvents as trimethylthiolignin. It commences to soften about 200°–205° C.

(e) *Tetracetyltrimethylthiolignin*

Trimethylthiolignin (20 g.) was acetylated under the conditions used for thiolignin. Yield 18.1 grams. After two purifications from dioxan-ether, yield 17.4 grams.

Found: C, 62.1; H, 5.9; S, 2.0;  $\text{OCH}_3$ , 24.9;  $\text{COCH}_3$ , 11.0%.

Calculated for  $\text{C}_{79}\text{H}_{90}\text{O}_{29}\text{S}$ : C, 61.7; H, 5.9; S, 2.1;  $\text{OCH}_3$ , 24.2;  $\text{COCH}_3$ , 11.2%.

Tetracetyltrimethylthiolignin is a brownish powder which commences to soften at 165°–170° C. It is soluble in the same solvents as trimethylthiolignin.

(f) *Hexamethylthiolignin*

(i) *Method A.*—Tetracetyltrimethylthiolignin (15 g.) was dissolved in acetone (200 cc.) and warmed to 40° C. Dimethyl sulphate (150 cc.) and 30 per cent. sodium hydroxide (400 cc.) were then added in small portions over a period of three hours, the temperature being maintained at 40°–45° C. by the application of an ice bath when necessary, and the mixture being vigorously stirred throughout. Stirring was continued for a further two hours, after which the acetone was removed under reduced pressure. The crude product separated out as a tar which solidified overnight. This was filtered off, ground up in a mortar with water, filtered, and washed with water. Yield 10.9 grams. A sample was purified by dissolving in ten volumes anhydrous dioxan, filtering and precipitating from anhydrous ether. Found:  $\text{OCH}_3$ , 31.0%. The product was remethylated twice under the same conditions. Yield 9.0 grams.

Found: C, 63.2; H, 6.3; S, 2.2;  $\text{OCH}_3$ , 33.0%; mol. wt. (Rast), 1416.

Calculated for  $\text{C}_{74}\text{H}_{88}\text{O}_{25}\text{S}$  (1408): C, 63.2; H, 6.3; S, 2.3;  $\text{OCH}_3$ , 33.0%; mol. wt., 1408.

Attempts to increase the methoxyl by a further remethylation yielded a product with 26.7 per cent.  $\text{OCH}_3$ .

Hexamethylthiolignin is a dull yellow powder, soluble in the same solvents as trimethylthiolignin, and commences to soften at 220°–225° C.

(ii) *Method B.*—Trimethylthiolignin (9 g.) was dissolved in anhydrous acetone (90 cc.) and redistilled dimethyl sulphate (15 cc.), and anhydrous potassium carbonate (15 g.) added. The mixture was refluxed gently for 24 hours, filtered, and the residue washed with dry acetone. The combined filtrate and washings were added dropwise to dry ether (1,000 cc.) with stirring, and the product filtered and washed with dry ether. Yield 6.7 grams;  $\text{OCH}_3$

(after two purifications from dioxan-ether) 33.2 per cent. Remethylation yielded a product with 29.4 per cent.  $\text{OCH}_3$ .

(g) *Acetylhexamethylthiolignin*

Hexamethylthiolignin (3.1 g.), made by Method A, was dissolved in anhydrous pyridine (15.5 cc.) and acetic anhydride (11 cc.) added. The mixture was allowed to stand 48 hours at room temperature and then added dropwise to vigorously stirred dry ether (300 cc.). The precipitate was filtered and washed with dry ether. Yield 3.0 grams. After two purifications from dioxan-ether, yield 2.7 grams.

Found: C, 63.0; H, 6.2; S, 2.1;  $\text{OCH}_3$ , 32.2;  $\text{COCH}_3$ , 2.9%.

Calculated for  $\text{C}_{76}\text{H}_{90}\text{O}_{26}\text{S}$ : C, 62.9; H, 6.2; S, 2.2;  $\text{OCH}_3$ , 32.1;  $\text{COCH}_3$ , 3.0%.

Acetylhexamethylthiolignin is a dull yellow powder soluble in the same solvents as trimethylthiolignin. It commences to soften at  $185^\circ\text{C}$ .

#### IV. ACKNOWLEDGMENTS

The authors are indebted to Messrs. Australian Paper Manufacturers Limited for permission to publish this work; and to Miss Betty Hickox and Mr. N. Lottkowitz who carried out many of the analyses.

#### V. REFERENCES

- (1) KLASON, P., and SEGERFELD, B.—*Ark. Kemi Min. Geol.* **4**(6), (1913).
- (2) ANDERZEN, O., and HOLMBERG, B.—*Ber. dtsh. chem. Ges.* **56B**: 2,044 (1923).
- (3) CINVES, F. J.—*Paper Tr. J.* **91**(19): 56 (1930).
- (4) HÄGGLUND, E.—*Svensk PappTidn.* **44**(9): 183 (1941).
- (5) HÄGGLUND, E., and LUMME, K.—*Papp. Travarutidsk. Finland* **5**: 361 (1925).
- (6) HANSON, F. S.—*Paper Tr. J.* **112**(2): 32 (1941).
- (7) KIMBLE, G. K.—*Tech. Ass. Pap. (TAPPI)* **25**: 341 (1942).
- (8) KULLGREN, C.—*IngenVetenskAkad. Handl.* No. 65, 35 pp. (1927). (*Chem. Abstr.* **22**: 1,680 (1928).)
- (9) BRAUNS, F. E.—*Paper Tr. J.* **111**(14): 33 (1940).
- (10) AHLM, C. E.—*Ibid.* **113**(13): 115 (1941).
- (11) COOKE, R. G.—*Aust. Pat. App.* No. 18,819. (Jan. 25, 1945).
- (12) FEIGL, F.—“Laboratory Manual of Spot Tests”, p. 180. (Academic Press: New York, 1943.)
- (13) FREUDENBERG, K., and SOHNS, F.—*Ber. dtsh. chem. Ges.* **66B**: 262 (1933).  
FREUDENBERG, K., SOHNS, F., and JANSON, A.—*Liebigs Ann.* **518**: 62 (1935).





# THE THEORY OF SYMMETRICAL CROSSED FLEXURE PIVOTS

By W. H. WITTRICK\*

(Plate 1)

[*Manuscript received February 17, 1948*]

## *Summary*

The general theory of symmetrical crossed flexure pivots under any combination of vertical and horizontal loads is given. This is used to derive a set of curves presenting all the information necessary for the design of such pivots and a set of experiments verifying the theory is described.

## I. INTRODUCTION

Crossed flexure pivots of the type shown in Figure 1 are in common use in many scientific instruments, particularly in balances where they are used in place of knife edge bearings. They have many qualities in their favour, the main ones being that they are not susceptible to wear, they do not require lubrication, they do not suffer from hysteresis or backlash and are of very simple and robust construction. They can also withstand sudden changes of load which may unseat a knife edge. Their main disadvantage lies in the fact that they are not suitable for very large angles of rotation.

The most comprehensive treatment of this subject appears to be that of Eastman(1, 2). These two papers give most of the basic information necessary for the design of flexure pivots but it was felt that there was need for a further and more searching investigation, particularly into the nature of the assumptions made and the effect of side loads on pivots. Also it should be possible to present the results of the theory more concisely. The analysis set out in this paper is an attempt to achieve this object.

Partly for the sake of simplicity and partly to conform with general practice, attention was concentrated on pivots in which the flexure strips are symmetrical with respect to the vertical centre line and cross at their mid-points. This means that the strips on each side of the pivot have the same thickness, length and total width, as indicated in Figure 1. A further limitation imposed on the analysis is the fact that only small rotations (up to about  $20^\circ$ ) are considered but it is felt that this will cover the great majority of practical cases.

\* Department of Aeronautical Engineering, University of Sydney.

In Section II the general theory is given and this is applied in Section III to the case of a  $45^\circ$  flexure pivot in which the strips are at right angles to each other. Design curves are given in Figures 7, 9, and 10 which enable the rotational stiffness, strip tensions and maximum bending moments to be quickly calculated for any combination of vertical and horizontal loads carried by the pivot. It is pointed out that these curves, although obtained primarily for a  $45^\circ$  pivot, are applicable, by a simple change of coordinates, to the case of a pivot with the strips inclined at any angle to the vertical.

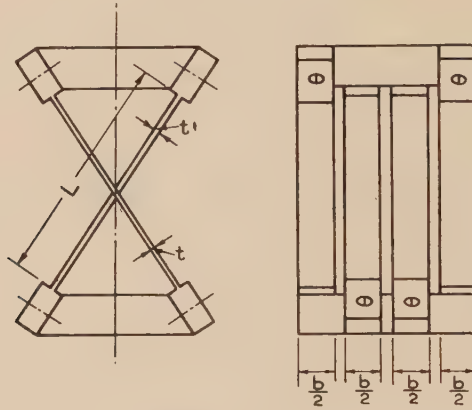


Fig. 1.—Arrangement of typical flexure pivot.

It is shown that this type of pivot is stable only over a limited range of vertical and horizontal loads (Fig. 8). Instability may occur with the vertical load tending to induce either tension or compression in the strips, or with a horizontal load acting alone. This phenomenon was further investigated experimentally using a large scale pivot. The results agreed very closely with the theoretical behaviour and a short description of the experiment is given in Section IV.

#### Notation

$L$  = length of each flexure strip.

$b$  = total width of flexure strips on each side of the pivot.

$t$  = thickness of each flexure strip.

$\alpha$  = inclination of each flexure strip to the vertical.

$I = bt^3/12$  = relevant moment of inertia of the cross section of the flexure strips on each side of the pivot.

$E$  = Young's modulus of the material.

$V$  = vertical load carried by the moving block (suffixes  $T$  and  $C$  indicate that the force is tending to induce tension and compression respectively in the strips).

$H$  = horizontal load carried by the moving block.

$C$  = moment about the pivot centre of the forces acting on the moving block.

$\theta$  = angle of rotation of the moving block in the direction of  $C$ .

$P$  = total direct force in the strips on one side of the pivot (suffixes  $T$  and  $C$  indicate that the strips are in tension and compression respectively).

$F$  = total shear force in the strips on one side of the pivot.

$$\beta_T = \sqrt{\frac{P_T L^2}{4EI}}$$

$$\beta_C = \sqrt{\frac{P_C L^2}{4EI}}$$

$M$  = total bending moment in the strips on one side of the pivot.

$$\varphi = \beta_T (\coth \beta_T - \beta_T)$$

$$\psi = \beta_C (\cot \beta_C + \beta_C)$$

Suffixes 1 and 2 refer to the strips on the two sides of the pivot.

## II. GENERAL THEORY

### (a) *The Centre of Rotation*

Suppose that under some system of forces and moments applied to the moving block the pivot is displaced as shown in Figure 2. The point  $A_1$  moves to  $A'_1$  through distances  $\delta_1$ , in a direction perpendicular to  $A_1 B_1$ , and  $e_1$ , in a direction parallel to  $A_1 B_1$ . The point  $A_2$  moves similarly to  $A'_2$  through distances  $\delta_2$  and  $e_2$ .

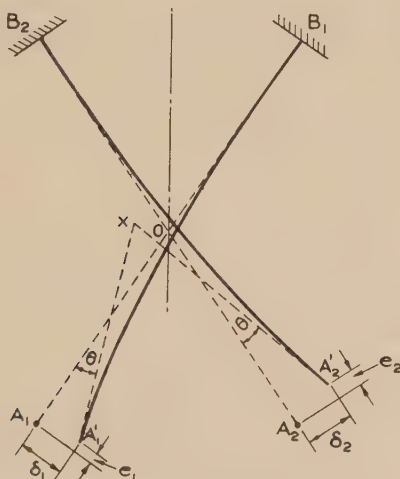


Fig. 2.—Definition of displacements.

The “transverse” displacement  $\delta$  is caused entirely by the bending of the strip. The “shortening”  $e$  is due partly to the change of slope and partly to the direct load (tension or compression) in the strip.

If the moving block is rigid then  $A'_1A'_2=A_1A_2$

$$\therefore (\delta_2 - \delta_1) = (e_1 + e_2) \tan \alpha.$$

Also, if the rotation of the moving block is  $\theta$ ,

$$(\delta_1 + \delta_2) = \theta L + (e_1 - e_2) \cot \alpha.$$

On solving for  $\delta_1$  and  $\delta_2$  these two equations give

$$\left. \begin{aligned} \delta_1 &= \frac{1}{2}\theta L + e_1 \cot 2\alpha - e_2 \operatorname{cosec} 2\alpha \\ \delta_2 &= \frac{1}{2}\theta L + e_1 \operatorname{cosec} 2\alpha - e_2 \cot 2\alpha. \end{aligned} \right\}$$

In the majority of cases the angle  $\alpha = 45^\circ$  approximately so that

$$\begin{aligned} \delta_1 &= \frac{1}{2}\theta L - e_2 \\ \delta_2 &= \frac{1}{2}\theta L + e_1. \end{aligned}$$

Further the shortening  $e$  is usually very small in comparison with the transverse deflection  $\delta$  so that finally we have

$$\delta_1 = \delta_2 = \frac{1}{2}\theta L \text{ very nearly} \dots\dots\dots (1)$$

This equation indicates that the intersection  $X$  of the tangents to the deflected strips at  $A_1$  and  $A_2$  coincides with  $O$  which has a constant position relative to the fixed block. Hence the centre of rotation of the pivot remains very nearly fixed and can only move appreciably if the displacement  $e$  can no longer be neglected in comparison with  $\delta$ . The problem is considerably simplified if we can fix the centre of rotation, and henceforth it will be assumed that equation (1) is valid.

### (b) The Bending of a Single Strip

(i) *If the Strip is in Tension.*—Consider the case of a single strip  $AB$  subjected to a system of end loads, as shown in Figure 3 (a), such that they cause a rotation and a deflection at  $A$  of  $\theta$  and  $\frac{1}{2}\theta L$  respectively. The differential equation for the deflection is

$$-EI \frac{d^2y}{dx^2} = P_T(\frac{1}{2}\theta L - y) - F_T x - M_A.$$

$$\text{Let } \beta_T^2 = \frac{P_T L^2}{4EI} \dots\dots\dots (2)$$

$$\therefore \frac{d^2y}{dx^2} - \left(\frac{2\beta_T}{L}\right)^2 y = \left(\frac{2\beta_T}{L}\right)^2 \left(\frac{F_T}{P_T} x + \frac{M_A}{P_T} - \frac{1}{2}\theta L\right).$$

The solution of this equation is

$$y = R \sinh \frac{2\beta_T x}{L} + S \cosh \frac{2\beta_T x}{L} - \frac{F_T}{P_T} x - \frac{M_A}{P_T} + \frac{1}{2}\theta L \dots\dots\dots (3)$$

where  $R$  and  $S$  are constants of integration.

The boundary conditions are

$$(i) \text{ at } x=0, \frac{dy}{dx} = -\theta, y = \frac{1}{2}\theta L.$$

$$(ii) \text{ at } x=L, y = \frac{dy}{dx} = 0.$$



On substituting these conditions in (3) and simplifying the resulting expressions we obtain

$$R = -\frac{L\theta}{4\beta_T} \dots\dots\dots (4)$$

$$S = \frac{L\theta}{4\beta_T} \coth \beta_T \dots\dots\dots (5)$$

$$M_A = \frac{EI\theta}{L} \beta_T \coth \beta_T \dots\dots\dots (6)$$

$$F_T = \frac{1}{2}\theta P_T \dots\dots\dots (7)$$

This last equation shows that the resultant reaction at the end *A* passes through *B*. Hence

$$M_B = M_A \dots\dots\dots (8)$$

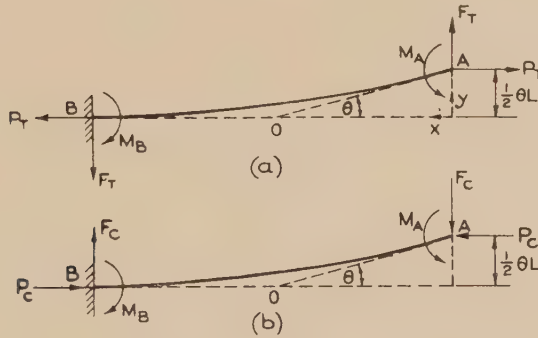


Fig. 3.—Definition of actions on a single strip.

The bending moment *M* at any point in the strip is equal to  $-EI(d^2y/dx^2)$  which, from (3) in conjunction with (4) and (5), becomes

$$M = \frac{EI\theta}{L} \beta_T \frac{\cosh \beta_T \left(1 - \frac{2x}{L}\right)}{\sinh \beta_T} \dots\dots\dots (9)$$

It will be seen from (9) that the maximum bending moment  $M_{max}$  occurs at the ends of the strip so that

$$M_{max} = M_A = \frac{EI\theta}{L} \beta_T \coth \beta_T \dots\dots\dots (10)$$

The moment about the point *O* of the external forces applied at *A* is given by

$$M_O = M_A - \frac{1}{2}\theta LP_T + \frac{1}{2}LF_T.$$

Using (2), (6), and (7) this becomes

$$M_O = \frac{EI\theta}{L} \cdot \varphi \dots\dots\dots (11)$$

$$\text{where } \varphi = \beta_T (\coth \beta_T - \beta_T) \dots\dots\dots (12)$$

The function  $\varphi$  is shown plotted against  $\beta_T$  in Figure 4.

(ii) *If the Strip is in Compression.*—Suppose that the end loads acting on the strip are as shown in Figure 3 (b). Then the equations corresponding to those derived for the tension case can be obtained directly by the substitutions

$$\begin{aligned} P_C &= -P_T \\ \beta_C &= \sqrt{-1} \beta_T \\ F_C &= -F_T \end{aligned}$$

$$\text{where } \beta_C^2 = \frac{P_C L^2}{4EI} \dots\dots\dots (13)$$

$$\text{Thus } M_A = \frac{EI\theta}{L} \beta_C \cot \beta_C \dots\dots\dots (14)$$

$$F_C = \frac{1}{2} P_C \dots\dots\dots (15)$$

$$M_B = M_A \dots\dots\dots (16)$$

$$\text{and } M = \frac{EI\theta}{L} \beta_C \frac{\cos \beta_C \left(1 - \frac{2x}{L}\right)}{\sin \beta_C} \dots\dots\dots (17)$$

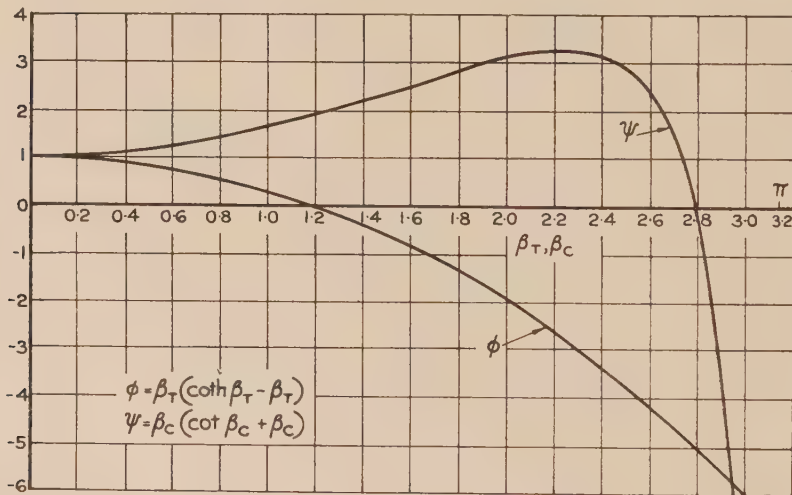


Fig. 4.—Moment coefficients for strips in tension and compression.

From (17) we see that in this case the maximum bending moment  $M_{max}$  occurs at the centre of the strip so that

$$M_{max} = \frac{EI\theta}{L} \beta_C \operatorname{cosec} \beta_C = M_A \sec \beta_C \dots\dots\dots (18)$$

Also, as  $\beta_C \rightarrow \pi$ ,  $M \rightarrow \infty$  and buckling occurs.

The moment about the point  $O$  of the external forces applied at  $A$  is given by

$$M_O = \frac{EI\theta}{L} \cdot \psi \dots\dots\dots (19)$$

$$\text{where } \psi = \beta_C (\cot \beta_C + \beta_C) \dots\dots\dots (20)$$

The function  $\psi$  is shown plotted against  $\beta_C$  in Figure 4.

(c) *The Forces in the Strips*

Having obtained the fundamental equations for the bending of a single strip we are now in a position to consider the pivot as a whole. Suppose that the moving block is subjected to a vertical force  $V$  and a side force  $H$  applied through  $O$ , the centre of the pivot, together with a couple  $C$  as shown in Figure 5. Let the reactions exerted by the strips on the moving block be  $P$ ,  $F$ , and  $M_A$  (with suffixes 1 and 2 denoting a particular strip). The force  $P$  is counted positive when the strip is in tension.

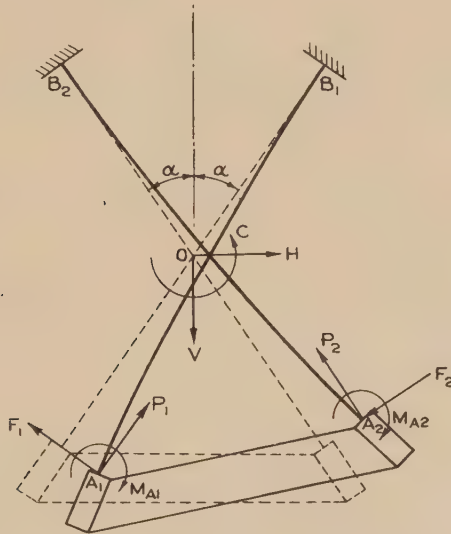


Fig. 5.—Forces and moments on displaced pivot.

Resolving vertically and horizontally the following equations are obtained :

$$\left. \begin{aligned} V &= (P_1 + P_2) \cos \alpha + (F_1 - F_2) \sin \alpha \\ H &= (F_1 + F_2) \cos \alpha - (P_1 - P_2) \sin \alpha \end{aligned} \right\} \dots\dots\dots (21)$$

Now whether the strips are in tension or compression we have from (7) and (15)

$$F = \frac{1}{2} \theta P.$$

Substituting this in (21) and solving for  $P_1$  and  $P_2$  gives

$$\left. \begin{aligned} P_1 &= \frac{2}{\theta} F_1 = \frac{V(2 \tan \alpha + \theta) - H(2 - \theta \tan \alpha)}{(4 + \theta^2) \sin \alpha} \\ P_2 &= \frac{2}{\theta} F_2 = \frac{V(2 \tan \alpha - \theta) + H(2 + \theta \tan \alpha)}{(4 + \theta^2) \sin \alpha} \end{aligned} \right\} \dots\dots\dots (22)$$

In the great majority of cases  $\theta$  is small enough to be neglected in (22). This amounts to neglecting the shear forces in the strips in comparison with the tensions in the equations of equilibrium (21). With this approximation we have

$$\left. \begin{aligned} P_1 &= \frac{2}{\theta} F_1 = \frac{1}{2}(V \sec \alpha - H \operatorname{cosec} \alpha) \\ P_2 &= \frac{2}{\theta} F_2 = \frac{1}{2}(V \sec \alpha + H \operatorname{cosec} \alpha) \end{aligned} \right\} \dots\dots\dots (23)$$

From these two equations it will be seen that if  $|H| < |V| \tan \alpha$  both strips are in tension or both in compression according to the direction of  $V$ . However, if  $|H| > |V| \tan \alpha$  one strip is in tension and one in compression.

#### (d) *The Rotational Stiffness of the Pivot*

Reverting to Figure 5, the equation of moment equilibrium for the moving block is

$$C = M_{01} + M_{02}$$

where  $M_0$  is defined by either (11) or (19), according to whether the strips are in tension or compression.

Hence

$$\frac{CL}{EI\theta} = (\varphi_1 \text{ or } \psi_1) + (\varphi_2 \text{ or } \psi_2)$$

the choice depending on the direct load in the strip.

It will be seen from (22) that the values of  $P_1$  and  $P_2$ , and therefore  $\beta_1$  and  $\beta_2$ , are not strictly independent of  $\theta$ . Hence the applied couple  $C$  is not exactly proportional to the rotation  $\theta$ . However, if the approximate equations (23) are used this complication is overcome and we may define a stiffness  $K = C/\theta$  which is dependent only on the properties of the pivot and upon the direct loads  $V$  and  $H$  applied. The error involved in making this approximation is small.

$$\text{Thus } KL/EI = (\varphi_1 \text{ or } \psi_1) + (\varphi_2 \text{ or } \psi_2) \dots\dots\dots (24)$$

### III. 45° FLEXURE PIVOT

In this section the preceding results are used to derive curves giving the stiffness, strip tensions and maximum bending moment for given applied loads  $V$  and  $H$  in the case of a pivot in which the strips are at right angles to each other.

From equations (23) we have, with  $\alpha = 45^\circ$ ,

$$\left. \begin{aligned} \frac{P_1 L^2}{EI} &= \frac{1}{\sqrt{2}} \left( \frac{VL^2}{EI} - \frac{HL^2}{EI} \right) \\ \frac{P_2 L^2}{EI} &= \frac{1}{\sqrt{2}} \left( \frac{VL^2}{EI} + \frac{HL^2}{EI} \right) \end{aligned} \right\} \dots\dots\dots (25)$$



These equations, in conjunction with (2) and (13), enable us to construct a chart showing the values of  $\beta_T$  or  $\beta_C$  for each of the two strips for a series of values of  $VL^2/EI$  and  $HL^2/EI$ . Since a strip buckles when the value of  $\beta_C$  becomes equal to  $\pi$ , the range of possible values of  $VL^2/EI$  and  $HL^2/EI$  is immediately limited to that shown in Figure 6, which also indicates whether the strips are in tension or compression. Along each of the dotted lines one or other of the strips has zero tension and along the cross-hatched boundary of the region one of the strips has buckled.

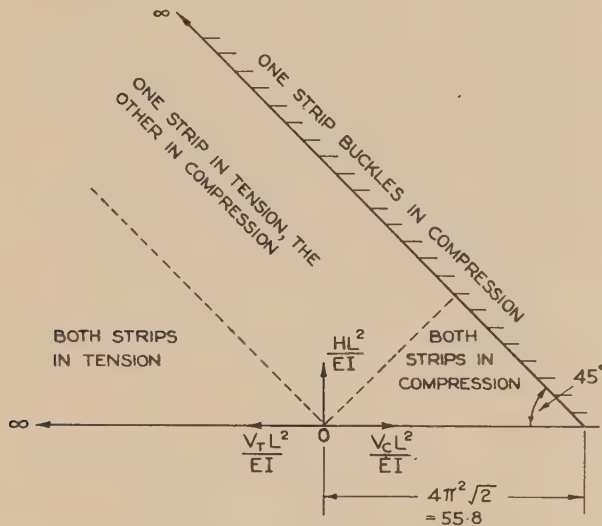


Fig. 6.—Boundaries of stability of the strips.

Knowing the values of  $\beta_T$  or  $\beta_C$  for each strip at a series of points in this region the corresponding values of  $\varphi$  or  $\psi$  can then be obtained from equations (12) and (20) (or from Fig. 4), and thence the value of  $KL/EI$  from equation (24). This has been done and the results are shown in Figure 7. The value of  $KL/EI$  is plotted against  $VL^2/EI$  and each curve represents a definite value of  $HL^2/EI$ . It is seen that there is only a limited region in which the pivot is stable, i.e. has a positive stiffness. Further, the addition of a side load to the pivot always makes it less stable, thus suggesting further bounds to the region of Figure 6. For the purposes of the results given here the region finally considered is as shown in Figure 8 which indicates also the extent of the range over which the pivot is stable. The boundary of this stable range gives the relation between  $VL^2/EI$  and  $HL^2/EI$  for pivots of zero stiffness.

With  $V=H=0$  the pivot stiffness  $K$  is equal to  $2(EI/L)$ , but by the introduction of a vertical compressive load  $V_C=30(EI/L^2)$  with no side load it is possible to increase the pivot stiffness to a maximum value of  $6.5(EI/L)$ .

The values of maximum bending moment in the strips can be obtained from Figure 9 which gives the value of  $(M_{max}L)/EI\theta$  plotted against  $VL^2/EI$  for each of the two strips 1 and 2. Each curve represents a definite value of  $HL^2/EI$ . These curves were obtained by means of the chart of values of  $\beta_T$  or  $\beta_C$  in conjunction with equations (10) and (18).

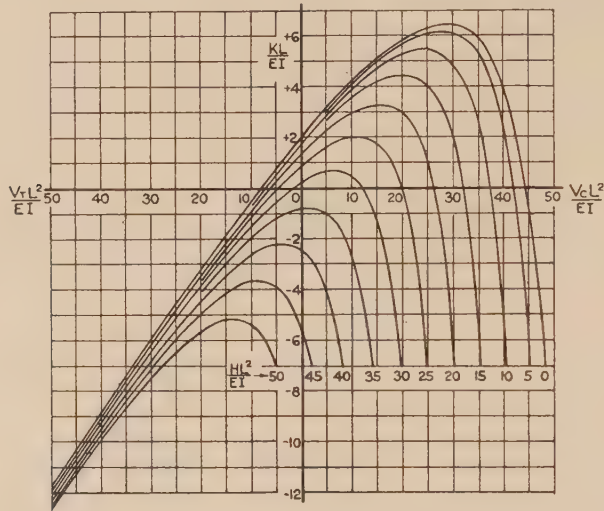


Fig. 7.—Stiffness diagram showing regions of positive and negative stability.

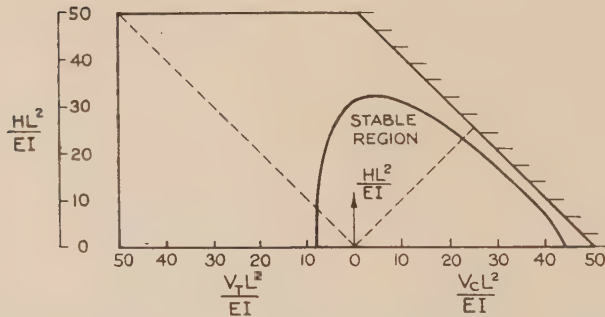


Fig. 8.—Extent of stable region and boundary for pivots of zero stiffness.

Figure 10 gives the corresponding tensions  $P$  in the strips, plotted as  $PL^2/EI$  against  $VL^2/EI$  for each of the two strips 1 and 2, and again each curve represents a definite value of  $HL^2/EI$ .

The three sets of curves in Figures 7, 9, and 10 give all the information necessary for the design of a  $45^\circ$  flexure pivot. However, it should be pointed

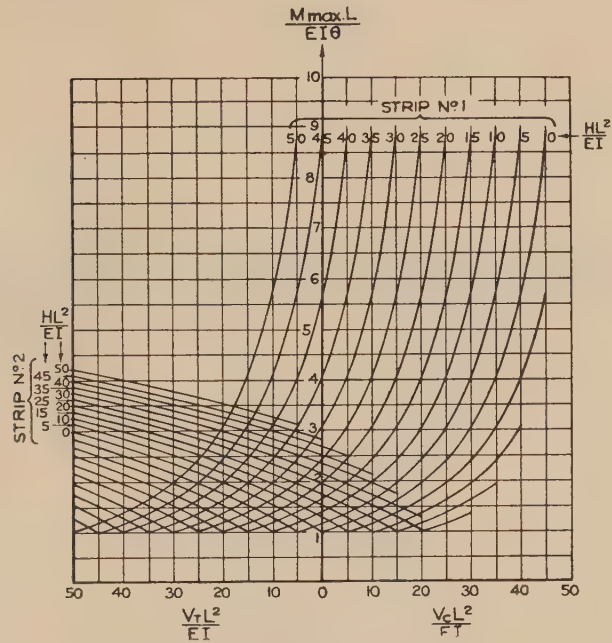


Fig. 9.—Curves for determining maximum bending moment in strips.

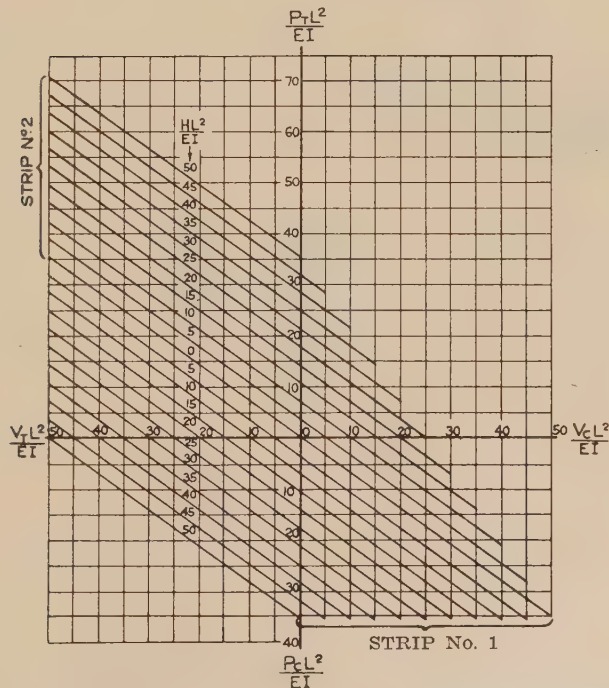


Fig. 10.—Curves for determining longitudinal force in strips.

out that the curves can be made to apply to a pivot of any angle  $\alpha$  simply by replacing  $VL^2/EI$  and  $HL^2/EI$  by  $\frac{VL^2}{EI} \frac{\sec \alpha}{\sqrt{2}}$  and  $\frac{HL^2}{EI} \frac{\operatorname{cosec} \alpha}{\sqrt{2}}$  respectively.

This will be readily seen from equations (23).

*Numerical Example.*—As an example of the use of Figures 7, 9, and 10 consider the case of a  $45^\circ$  flexure pivot with the following properties:  $L=0.5$  in.  $b=0.25$  in.,  $t=0.010$  in.,  $E=3 \times 10^7$  lb. per sq. in. The pivot carries a vertical compressive load  $V_C=40$  lb. and a side load  $H=25$  lb. It is required to find the stiffness  $K$  under these conditions and the maximum stresses in the strips when the angle of rotation is  $\theta=0.1$  radian ( $=5.73^\circ$ ).

From the above data  $\frac{EI}{L^2}=2.5$  lb.

$$\frac{EI}{L}=1.25 \text{ lb. in.}$$

Hence 
$$\frac{V_C L^2}{EI} = \frac{40}{2.5} = 16$$

and 
$$\frac{H L^2}{EI} = \frac{25}{2.5} = 10.$$

Then from Figure 7,

$$\frac{KL}{EI} = +4.88$$

$$\therefore K = +4.88 \times 1.25 = +6.10 \text{ lb. in. per radian.}$$

Also, from Figures 9 and 10 we have

$$\text{for strip 1 } \frac{M_{1\max} L}{EI\theta} = 2.55, \quad \frac{P_1 L^2}{EI} = 18.4 \text{ (compression)}$$

$$\text{and for strip 2 } \frac{M_{2\max} L}{EI\theta} = 1.20, \quad \frac{P_2 L^2}{EI} = 4.2 \text{ (compression).}$$

Thus  $M_{1\max} = 2.55 \times 1.25 \times 0.1 = 0.32$  lb. in.

$$P_1 = 18.4 \times 2.5 = 46.0 \text{ lb.}$$

$$M_{2\max} = 1.20 \times 1.25 \times 0.1 = 0.15 \text{ lb. in.}$$

$$P_2 = 4.2 \times 2.5 = 10.5 \text{ lb.}$$

Now the section modulus of each strip  $z = \frac{1}{6} \times 0.25 \times 0.01^2$   
 $= 1/240,000$  in.<sup>3</sup>

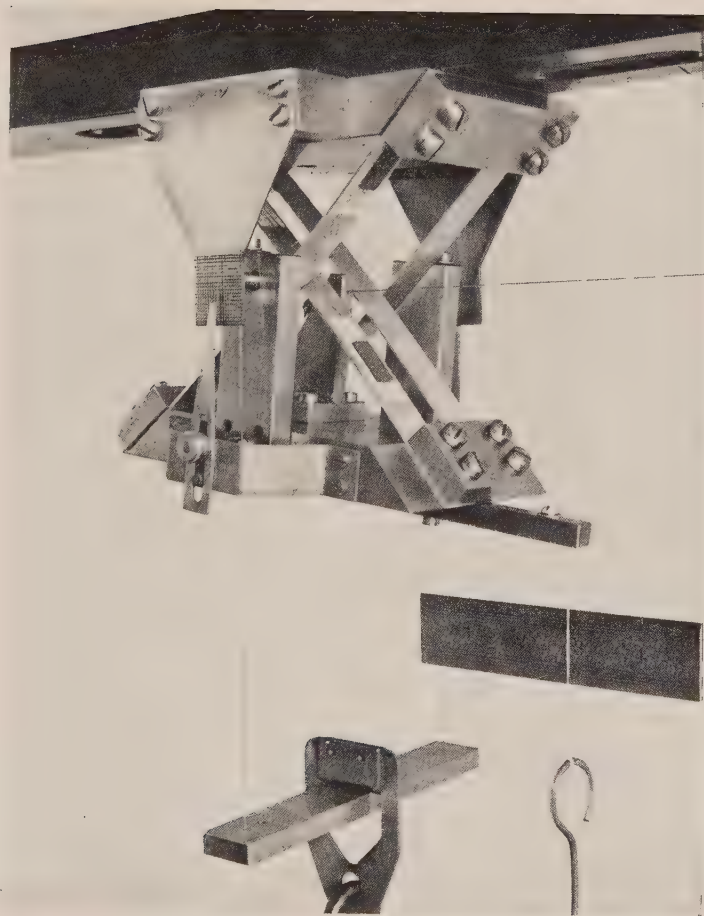
and the cross-sectional area of each strip  $a = 0.25 \times 0.01$   
 $= 1/400$  in.<sup>2</sup>

Hence the maximum stresses in the two strips are given by

$$f_{1\max} = 0.32 \times 240,000 + 46.0 \times 400 = 95,200 \text{ lb. per sq. in.}$$

$$\text{and } f_{2\max} = 0.15 \times 240,000 + 10.5 \times 400 = 40,200 \text{ lb. per sq. in.}$$





WITTRICK.—THE THEORY OF SYMMETRICAL CROSSED FLEXURE PIVOTS



## IV. EXPERIMENTAL INVESTIGATION

The results given in Figures 7, 9, and 10 exhibited so many interesting features that it was felt that an experimental check would be desirable. Consequently a large scale pivot was constructed having two duralumin strips on each side of the pivot. Each strip was 6 in. long, 0.75 in. wide (i.e.  $b=1.50$  in.), and 0.0478 in. thick. The pivot was designed so that it could be loaded by vertical and horizontal loads acting through the centre of the pivot and at the same time have a couple applied to it. An arrangement for watching the movement of the centre of rotation was also fitted. This took the form of a pointer attached to the moving block and a scribed scale attached to the fixed block. The pointer was set at the initial pivot centre and when rotation occurred its displacement over the scale gave a visual indication of the movement of the centre of rotation. A photograph of the pivot, set up for the application of a vertical tensile load together with a side load, is shown in Plate 1.

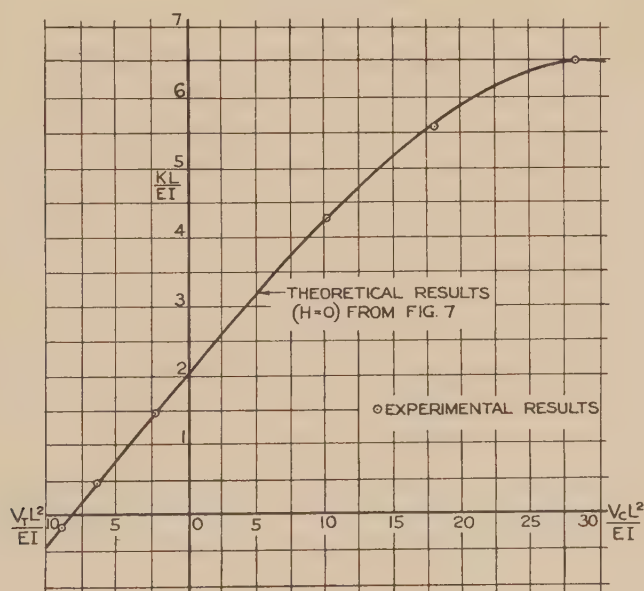


Fig. 11.—Comparison of experimental and theoretical results.

The first experiment consisted of investigating the variation of rotational stiffness with vertical load. The results, corrected for the restoring couple arising from the lateral movement of the centre of gravity of the moving block, are shown plotted in Figure 11. It will be seen that the agreement between experiment and theory is excellent. When a vertical tensile load greater than that required to cause instability (about 30 lb. in this case) was applied the pivot had to be restrained from rotating, thus demonstrating clearly the significance of a negative stiffness. It appears from Figure 11 that one of the experimental points was actually obtained with the pivot unstable. The explanation of this is that the restoring couple arising from the weight of the moving block was

slightly greater than that required to prevent the pivot from rotating so that in fact the pivot was still stable.

A further set of experiments designed to check the results for maximum bending moment given in Figure 9 was also carried out. The bending stresses were measured by means of resistance strain gauges cemented to the strips and again the agreement with theory was good. It was not as good as in the previous experiment but the errors were within the expected limit of accuracy of resistance strain gauges.

In all the experiments the greatest angle of rotation was about  $7^\circ$  and within this range the centre of rotation never moved by more than 0.05 in. Considering the size of the pivot this is remarkably small. However, it is probable that the movement of the centre of rotation will increase roughly as the square of the angle of rotation, so that for rotations as high as  $20^\circ$ , the movement may be considerably greater than indicated here.

## V. CONCLUSIONS

The main conclusions to be drawn from the results of the investigations described in this paper may be summarized as follows:

- (a) The centre of rotation of a crossed flexure pivot designed for small angles of rotation remains very nearly fixed.
- (b) The influence of direct loads on the properties of a pivot may be considerable.
- (c) There is only a limited range of direct loads which a pivot can withstand and still remain stable.
- (d) Theoretically, if the pivot carries direct loads, the angle of rotation is not exactly proportional to the applied couple. For all practical purposes, however, a linear relationship may be assumed.

## VI. ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance given in the experimental investigation by Mr. W. Howard, who designed and made the pivot, and by Mr. D. Barnsley, who carried out the experimental work.

## VII. REFERENCES

- (1) EASTMAN, F. S.—Flexure pivots to replace knife edges and ball bearings. Univ. Wash. Engng. Exp. Sta. Bull. No. 86 (1935).
- (2) EASTMAN, F. S.—The design of flexure pivots. *J. Aeronaut. Sci.* 5: 16-21 (Nov. 1937).

## EXPLANATION OF PLATE 1

Experimental crossed flexure pivot arranged for the simultaneous application of horizontal and vertical tensile loads.



# RELAXATION METHODS APPLIED TO TWO PROBLEMS OF TWO-DIMENSIONAL STRESS DISTRIBUTION INVOLVING MIXED BOUNDARY CONDITIONS

By W. H. WITTRICK\* and W. HOWARD\*

[Manuscript received January 27, 1948]

## Summary

Relaxation methods have been used to determine the stress distributions in both a rectangular and a highly tapered plate under tension when the load is applied through absolutely rigid clamps. Both problems require the treatment of boundary conditions involving the values of both stresses and displacements. The solutions were obtained in terms of displacements and the stresses were subsequently determined from them.

## I. INTRODUCTION

In the past, relaxation methods have been applied with success to the numerical solution of two types of two-dimensional stress problems(1, 2, 3). The first type is the determination of the stress distribution in a thin plate when the values of the tractions on all boundaries are completely specified. The usual procedure is to work in terms of the stress function for the distribution and the problem reduces to the determination of a function  $\phi$  such that  $\nabla^4\phi=0$  everywhere, and both  $\phi$  and  $\partial\phi/\partial n$  are known at all points on the boundary. Such a problem is quite amenable to treatment by a relaxation process but the  $\nabla^4$  operator is such that the convergence of the process is extremely slow, with the result that a considerable amount of time must be spent on the relaxation to obtain a solution.

The second type of problem is the determination of the stress distribution in a thin plate when the values of the displacements on all boundaries are completely specified. In this case the problem is reduced to the determination of the displacements  $u$  and  $v$  at points in the plate by the solution of two simultaneous linear partial differential equations of the second order (equations (2)) together with the specified boundary values of  $u$  and  $v$ . When solved by a relaxation process these equations are rapidly convergent in comparison with that of the first type of problem.

In the present paper relaxation methods have been used to determine the stress distribution in two problems which involve mixed boundary conditions, i.e. on some parts of the boundary the tractions are specified and on the remainder of the boundary the displacements are specified. The solution was obtained in terms of the displacements  $u$  and  $v$ , using the two simultaneous differential

\* Department of Aeronautical Engineering, University of Sydney.

equations mentioned above, but in this case the boundary conditions over those regions in which the stresses are specified, are considerably more complicated. However, the method adopted still gave the same rapid convergence.

The two problems chosen are quite simple in that all boundaries are straight. It was thought that the main features of the method could best be illustrated without the added complication of curved boundaries. However, there is no reason to suppose that the method cannot be extended to take account of these.

### Notation

- $x, y$  = rectangular coordinates of a point in a plate.  
 $u, v$  = component movements of the point  $(x, y)$ .  
 $E$  = Young's modulus of the material.  
 $\sigma$  = Poisson's ratio of the material.  
 $\widehat{xx}, \widehat{yy}, \widehat{xy}$  = stresses at the point  $(x, y)$ .  
 $h$  = length of the mesh square in the relaxation process.  
 $F_u, F_v$  = residuals in the relaxation process.

## II. THE FUNDAMENTAL DIFFERENTIAL EQUATIONS

### (a) Derivation

Let  $(u, v)$  be the movements of a point  $(x, y)$  in a plate subjected to a two-dimensional stress distribution.

The stresses at the point are given by

$$\left. \begin{aligned} \widehat{xx} &= \frac{E}{1-\sigma^2} \left\{ \frac{\partial u}{\partial x} + \sigma \frac{\partial v}{\partial y} \right\} \\ \widehat{yy} &= \frac{E}{1-\sigma^2} \left\{ \frac{\partial v}{\partial y} + \sigma \frac{\partial u}{\partial x} \right\} \\ \widehat{xy} &= \frac{E}{2(1+\sigma)} \left\{ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right\} \end{aligned} \right\} \dots\dots\dots (1)$$

The equations of equilibrium of an element in the absence of body forces are

$$\begin{aligned} \frac{\partial \widehat{xx}}{\partial x} + \frac{\partial \widehat{xy}}{\partial y} &= 0 \\ \frac{\partial \widehat{xy}}{\partial x} + \frac{\partial \widehat{yy}}{\partial y} &= 0. \end{aligned}$$

On substituting for  $\widehat{xx}, \widehat{yy}$  and  $\widehat{xy}$  in terms of  $u$  and  $v$ , these equations become

$$\left. \begin{aligned} 2 \frac{\partial^2 u}{\partial x^2} + (1-\sigma) \frac{\partial^2 u}{\partial y^2} + (1+\sigma) \frac{\partial^2 v}{\partial x \partial y} &= 0 \\ 2 \frac{\partial^2 v}{\partial y^2} + (1-\sigma) \frac{\partial^2 v}{\partial x^2} + (1+\sigma) \frac{\partial^2 u}{\partial x \partial y} &= 0 \end{aligned} \right\} \dots\dots\dots (2)$$

These fundamental equations in conjunction with appropriate boundary conditions are used to yield solutions for  $(u, v)$  in two particular problems to be described later.

(b) *Reduction of the Equations to a Relaxation Pattern*

The plate is considered to be divided into a square mesh of side  $h$ . The mesh-lines are parallel to the  $x$  and  $y$  directions. In terms of the calculus of finite differences, the first and second derivatives of  $u$  and  $v$  at a point 0 on the mesh become approximately

$$\left. \begin{aligned} \left( \frac{\partial u}{\partial x} \right)_0 &= \frac{1}{2h}(u_1 - u_3) \\ \left( \frac{\partial u}{\partial y} \right)_0 &= \frac{1}{2h}(u_2 - u_4) \\ \left( \frac{\partial^2 u}{\partial x^2} \right)_0 &= \frac{1}{h^2}(u_1 + u_3 - 2u_0) \\ \left( \frac{\partial^2 u}{\partial y^2} \right)_0 &= \frac{1}{h^2}(u_2 + u_4 - 2u_0) \\ \left( \frac{\partial^2 u}{\partial x \partial y} \right)_0 &= \frac{1}{4h^2}(u_a - u_b + u_c - u_d) \end{aligned} \right\} \dots\dots\dots (3)$$

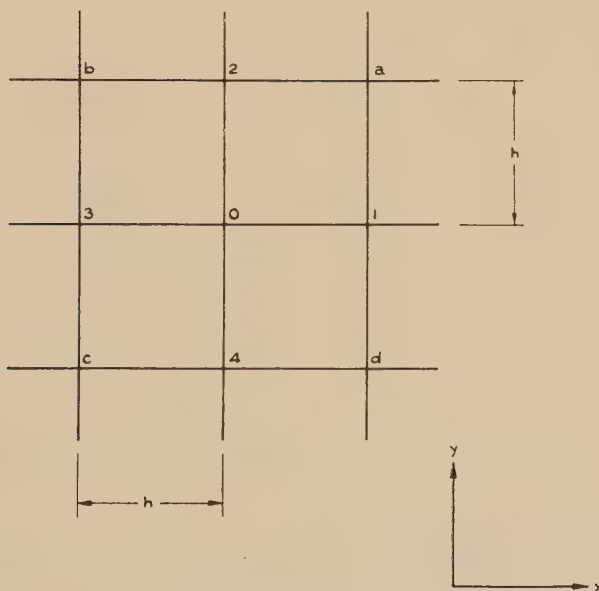


Fig. 1

where the suffixes refer to the point 0 and its surrounding mesh-points as shown in Figure 1. On substituting these values for the derivatives, equations (2) become

$$\left. \begin{aligned} 8(u_1 + u_3 - 2u_0) + 4(1 - \sigma)(u_2 + u_4 - 2u_0) + (1 + \sigma)(u_a - u_b + u_c - u_d) &= F_u \\ 8(v_2 + v_4 - 2v_0) + 4(1 - \sigma)(v_1 + v_3 - 2v_0) + (1 + \sigma)(u_a - u_b + u_c - u_d) &= F_v \end{aligned} \right\} \dots (4)$$

where  $F_u$  and  $F_v$  are the "residuals" for the relaxation process.

Approximately,  $\sigma=0.3$ , and henceforth this value will be used throughout the paper. With this substitution equations (4) become

$$\left. \begin{aligned} -21.6u_0 + 8(u_1 + u_3) + 2.8(u_2 + u_4) + 1.3(v_a - v_b + v_c - v_d) &= F_u \\ -21.6v_0 + 2.8(v_1 + v_3) + 8(v_2 + v_4) + 1.3(u_a - u_b + u_c - u_d) &= F_v \end{aligned} \right\} \dots (5)$$

From these equations, relaxation patterns can be derived which give the changes in the values of  $F_u$ ,  $F_v$  at adjacent mesh-points due to changes of  $(u, v)$  at the mesh-point 0. These patterns are given in Figure 2.

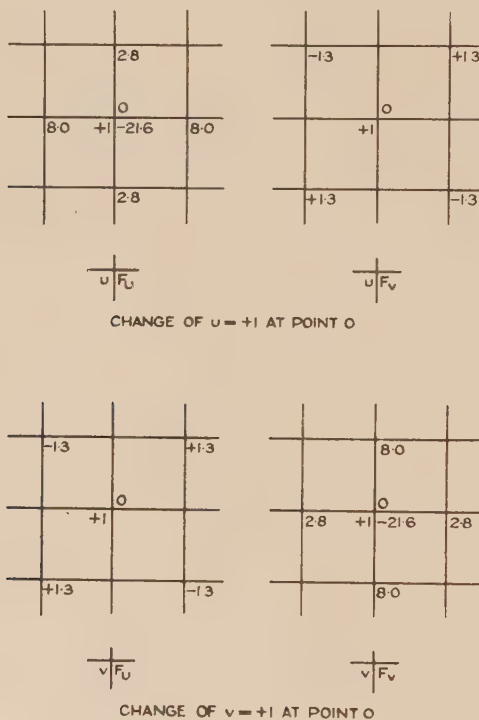


Fig. 2

### III. RECTANGULAR PLATE WITH FREE EDGES AND COMPLETELY CLAMPED ENDS UNDER TENSION

#### (a) Statement of the Problem

A rectangular plate of width  $b$  and length  $5b/2$  is subjected to a tensile force in the direction of its length (see Fig. 3). The force is considered to be applied by means of two absolutely rigid clamps in such a way that the ends of the plate are constrained to remain straight and parallel and the lateral contraction due to Poisson's ratio effect is prevented at the ends. The sides of the plate are quite free.



The general equations are

$$2 \frac{\partial^2 u}{\partial x^2} + 0.7 \frac{\partial^2 u}{\partial y^2} + 1.3 \frac{\partial^2 v}{\partial x \partial y} = 0$$

$$2 \frac{\partial^2 v}{\partial y^2} + 0.7 \frac{\partial^2 v}{\partial x^2} + 1.3 \frac{\partial^2 u}{\partial x \partial y} = 0.$$

Since these equations are linear no advantage is gained by using a non-dimensional notation, in fact  $u$ ,  $v$ ,  $x$ , and  $y$  may be treated as non-dimensional quantities.

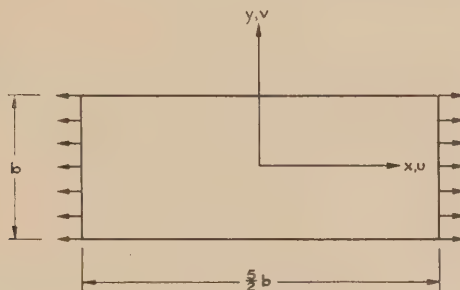


Fig. 3

The boundary conditions, referred to axes  $x$  and  $y$ , as shown in Figure 3, are

(i) at  $x = \frac{5b}{4}$ ,  $u = u_0$  (constant),  $v = 0$

(ii) at  $x = -\frac{5b}{4}$ ,  $u = -u_0$ ,  $v = 0$

(iii) at  $y = \pm \frac{b}{2}$ ,  $\widehat{yy} = \widehat{xy} = 0$

i.e.  $\frac{\partial v}{\partial y} + 0.3 \frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 0$ .

Since the plate is symmetrical about the  $x$  and  $y$  axes, only that quarter of the plate between  $x=0$ ,  $+5b/4$  and  $y=0$ ,  $+b/2$  will be considered.

#### (b) Treatment of Boundary Conditions in the Relaxation Process

The use of the relaxation process at mesh-points on, or adjacent to, the boundary needs special consideration. In addition to the general equations there are boundary conditions to be satisfied. It is obvious that the normal relaxation patterns cannot be used directly since the values of the residuals at points on the boundary involve the values of  $u$  and  $v$  at points outside the plate. Henceforth such points will be termed "image points" and the values of  $u$  and  $v$  at those points "image values".

(1) *Boundary  $x=5b/4$ .*—Since  $u$  and  $v$  are fixed along the boundary  $x=5b/4$ , no relaxation need be carried out on this boundary, and no account taken of residuals there.

(2) *Boundary*  $y=b/2$ .—On the boundary  $y=b/2$ , boundary conditions (iii) must be satisfied. Expressing these in terms of finite difference approximations at points 0, 1, 3 in turn (see Fig. 4) we have

$$\left. \begin{aligned} (v_2-v_4) + 0.3(u_1-u_3) &= 0 \\ (v_1-v_3) + (u_2-u_4) &= 0 \\ (v_a-v_d) + 0.3(u_I-u_0) &= 0 \\ (v_I-v_0) + (u_a-u_d) &= 0 \\ (v_b-v_c) + 0.3(u_0-u_{III}) &= 0 \\ (v_0-v_{III}) + (u_b-u_c) &= 0 \end{aligned} \right\} \dots\dots\dots (6)$$

From these equations, values of  $(u, v)$  at the image points  $b, 2, a$  are obtained in terms of  $(u, v)$  at mesh-points in the plate. Substitution of these image values in the expressions for  $F_u, F_v$  at the point 0 gives

$$\left. \begin{aligned} F_u &= -20.82u_0 + 5.6u_1 + 8(u_1+u_3) - 0.39(u_I+u_{III}) - 2.8(v_1-v_3) \\ F_v &= -19.0v_0 + 16v_1 + 2.8(v_1+v_3) - 1.3(v_I+v_{III}) - 2.4(u_1-u_3) \end{aligned} \right\} \dots (7)$$

From equations (7) the relaxation patterns for points on, and adjacent to the boundary  $y=b/2$  are obtained, as shown in Figures 5 and 6.

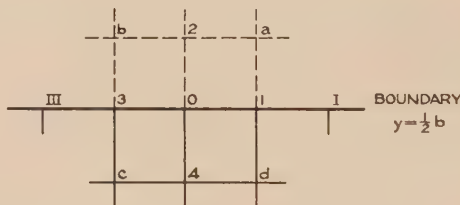


Fig. 4

It is clear from the above patterns that when  $u$  (or  $v$ ) is changed at a point on the boundary  $y=b/2$  the residual  $F_u$  (or  $F_v$ ) is affected to a distance of two mesh lengths along the boundary on either side of the point at which the change is made. Consequently when a change in  $u$  (or  $v$ ) is made on the boundary  $y=b/2$  at a point which is adjacent to the corner  $(5b/4, b/2)$ , the image values of  $u$  (or  $v$ ) in the boundary  $x=5b/4$  must be considered. These images are determined by the boundary conditions at the corner. Referring to Figure 7, since both  $u$  and  $v$  are constant along  $x=5b/4$ , then at the point 1,

$$\frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} = 0.$$

Also, since  $\widehat{yy} = \widehat{xy} = 0$  along  $y=b/2$ , then at the point 1 we have

$$\frac{\partial v}{\partial y} + \sigma \frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 0.$$

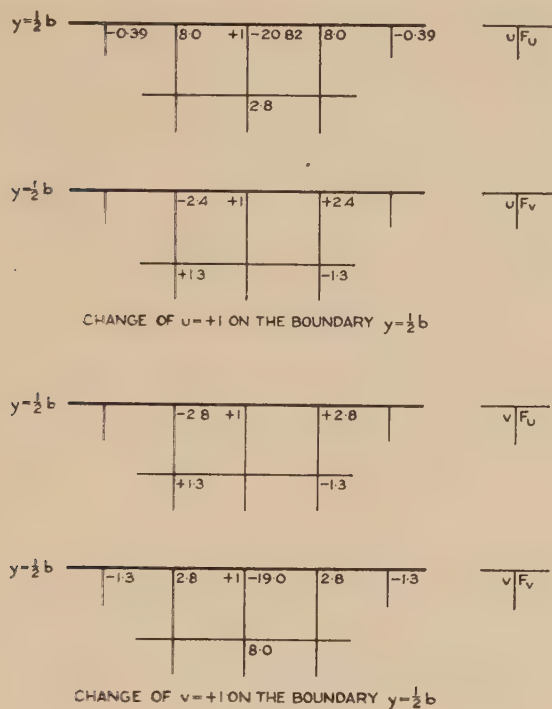


Fig. 5

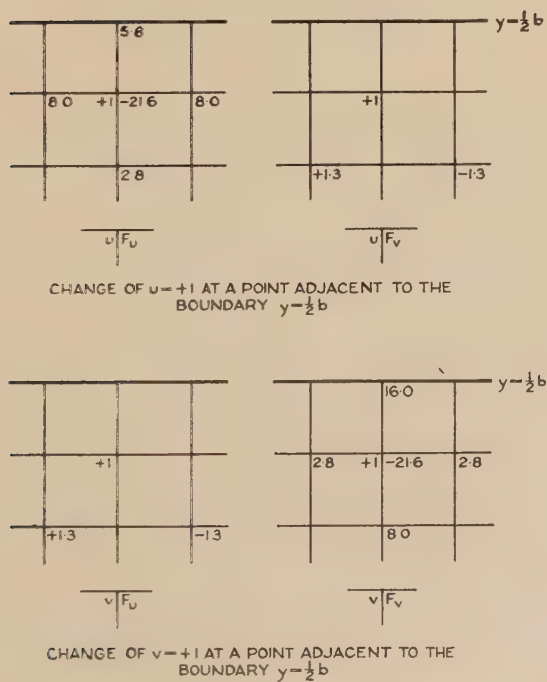


Fig. 6

Hence we see that at the point 1

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = \frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} = 0.$$

Applying the first and last of these equations, viz.  $\partial u/\partial x=0$ , and  $\partial v/\partial x=0$ , in terms of finite difference approximations, gives  $u_1=u_0$  and  $v_1=v_0$ . Consequently

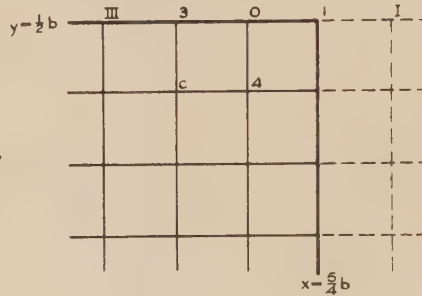


Fig. 7

if a change in  $u$  (or  $v$ ) is made at 0 (Fig. 7) then an equal change must be made in  $u$  (or  $v$ ) at the image point I and the special patterns for changes in  $u$  and  $v$  at the point 0 reduce to those given in Figure 8.

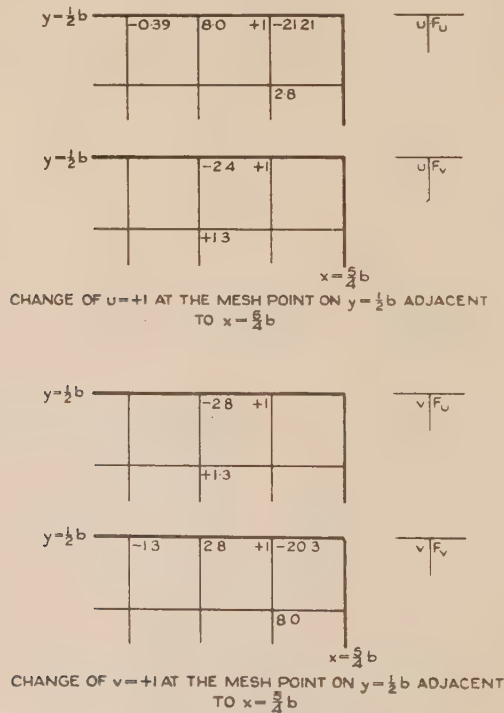


Fig. 8



(3) *Centre-lines*  $x=0, y=0$ .—When  $u$  and  $v$  are changed at points adjacent to the centre-lines of the plate, the appropriate changes in  $u$  and  $v$  must be made at the corresponding image points in the adjacent quarter-plate. This is best done mentally, as the process is so simple that special relaxation patterns are not necessary.

From the symmetry of the plate it is obvious that in the centre-line  $x=0$ ,  $u$  has an equal negative, and  $v$  an equal positive image, whilst in the centre-line  $y=0$ ,  $u$  has an equal positive, and  $v$  an equal negative image.

### (c) *Description of the Process*

The use of the relaxation technique, in obtaining the solution of linear partial differential equations, has already been fully described by other writers. It is not proposed to describe it further here. Only the salient features of this problem will be discussed.

A solution of the problem using a mesh-length of  $h=b/8$  was obtained before proceeding to a finer mesh.

A nominal value of 40 was taken for  $u_0$  (the displacement of the clamped end  $x=5b/4$ ). Once a solution of the distribution of  $(u, v)$  is obtained for one value of  $u_0$  then the solution for any other value follows by simple proportion.

First of all a guess was made at the distribution of  $u$  and  $v$  over the quarter-plate,  $u$  and  $v$  being taken to vary according to the equations

$$u=32x/b, \quad v=-32cy/b=-9.6y/b,$$

except that the value of  $v$  at the boundary  $x=5b/4$  was zero as required by boundary condition (i).

Using equations (3) and (5), and this distribution of  $(u, v)$ , residuals  $F_u, F_v$  were calculated at the mesh-points for  $h=b/8$ . With this distribution all the residuals were zero except at points where the values of  $v$  at the boundary  $x=5b/4$  appeared in the expressions for the residuals. This was because the assumed distribution of  $(u, v)$  in general satisfied the fundamental equations.

The residuals so calculated were "liquidated" using the relaxation patterns already derived. The residuals were reduced to such a value that a change of  $\pm 1$  in the last decimal place of either  $u$  or  $v$ , in general, caused an increase in the residuals. The process was very convergent and  $(u, v)$  were relaxed to the second decimal place for this mesh-length.

With the approximate solution for  $h=b/8$  obtained, values of  $(u, v)$  at intermediate points for the  $h=b/16$  mesh were derived by interpolation. In regions where the second derivatives of  $u$  and  $v$  were not small a parabolic interpolation formula was adopted, but in other regions it was found sufficient to use a straight line interpolation. In any case, inaccuracies in the interpolation are included in the magnitudes of the residuals on the finer mesh. The better the interpolation the smaller the residuals and hence the smaller the amount of work required to perform the relaxation.

Residuals were again calculated and reduced by the relaxation process, this time to such a value that a change of  $\pm 0.002$  in either  $u$  or  $v$ , in general, caused an increase in the residuals.

It was considered desirable to use a still finer mesh near the corner ( $5b/4, b/2$ ) since the solution was exhibiting a stress concentration there, but at the same time it was not considered necessary to extend the relaxation over the whole quarter-plate. Accordingly  $(u, v)$  along the lines  $x=15b/16$  and  $y=b/4$  were fixed at the values obtained by the  $h=b/16$  relaxation, and residuals derived for  $h=b/32$  mesh-points in the corner of the plate bounded by those two lines. The values of  $(u, v)$  at the intermediate points were again obtained by interpolation. This time, however, the interpolated values along the lines  $x=15b/16, y=b/4$  were important since these values remained fixed and occurred as terms in the expressions for  $F_u, F_v$  at the adjacent mesh-points. Thus care had to be exercised in this interpolation.

This measure of limiting the process to a certain region could only be adopted on the assumption that  $(u, v)$ , or rather their first derivatives, would not alter considerably near the edge of that region during the subsequent relaxation. This was found to be the case. Values of the predominant stress ( $\widehat{xx}$ ) calculated at mesh-points a distance of  $b/16$  inside the corner relaxed, differed by a negligible amount from the values obtained at the same points with the  $h=b/16$  mesh.

#### (d) Derivation of Stresses

Using the values of  $(u, v)$  found by the relaxation process, the stresses  $\widehat{xx}$ ,  $\widehat{yy}$ , and  $\widehat{xy}$  were obtained at the mesh-points from equations (1). The derivatives of  $u$  and  $v$  required for this computation were obtained from the finite difference approximations given in equations (3). For presentation of the results each of the stresses  $\widehat{xx}$ ,  $\widehat{yy}$ , and  $\widehat{xy}$ , is given in the non-dimensional form

$$(1-\sigma^2) \cdot b/u_0 \cdot \text{stress}/E.$$

The results of the  $h=b/16$  mesh were used throughout except in the corner bounded by  $x=b, y=9b/16$ , where the results of the  $h=b/32$  mesh were used. The stresses at all mesh-points within the boundaries of the plate were thus calculated.

The stresses on the boundaries needed special consideration since, in general, their computation involved image values of  $u$  and  $v$ .

(1) *Boundary  $y=b/2$ .*—The boundary condition requires that

$$\widehat{yy} = \widehat{xy} = 0.$$

Hence 
$$\frac{\partial v}{\partial y} = -\sigma \frac{\partial u}{\partial x},$$
 and the expression for  $\widehat{xx}$  becomes

$$(1-\sigma^2) \frac{b}{u_0} \frac{\widehat{xx}}{E} = \frac{b}{u_0} (1-\sigma^2) \frac{\partial u}{\partial x}.$$

(2) *Boundary  $x=5b/4$ .*—The values of  $\partial u/\partial x$  and  $\partial v/\partial x$  on the boundary  $x=5b/4$  were not obtainable directly from the values of  $u$  and  $v$  at points in the plate, since the finite difference approximations involve the values of  $(u, v)$  at the image points in the boundary  $x=5b/4$ . These image values were obtained from the fact that the two fundamental differential equations (2) must be satisfied on the boundary. This led to the formulation of a number of simultaneous equations in the image values of  $u$  and  $v$  which were solved by the use of a subsidiary relaxation process. The image values at the point  $\left(\frac{5b}{4}, \frac{b}{2}\right)$  were known (see (2) of Section III (b)).

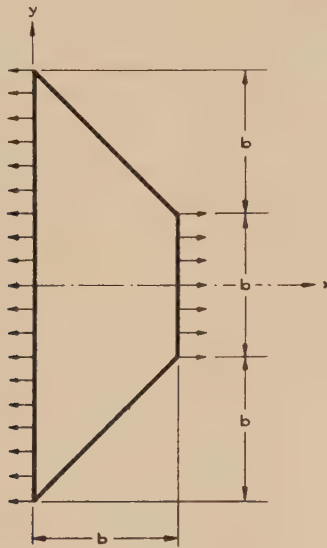


Fig. 9

#### IV. 45° TAPERED PLATE WITH FREE EDGES AND COMPLETELY CLAMPED ENDS UNDER TENSION

##### (a) *Statement of the Problem*

A straight tapered plate of length  $b$ , and width varying between  $b$  and  $3b$ , at its ends has one axis of symmetry, as shown in Figure 9. It is subjected to an axial tensile force in the direction of its length. As in the previous problem, the force is considered to be applied by means of two absolutely rigid clamps in such a way that the ends of the plate are constrained to remain straight and parallel and the lateral contraction due to Poisson's ratio effect is prevented at the ends. The tapered edges of the plate are quite free.

Axes are taken as shown in Figure 9.

The boundary conditions are

- (i) at  $x=0$ ,  $u=0$ ,  $v=0$
- (ii) at  $x=b$ ,  $u=u_0$  (constant),  $v=0$
- (iii) along the free edges given by  $x \pm y = 3b/2$ 
  - (a) the normal stress  $\widehat{nn}=0$ .
  - (b) the shearing stress  $\widehat{ns}=0$ .

The plate is symmetrical about the  $x$ -axis, hence only that half of the plate for which  $y$  is positive will be considered.

(b) *Treatment of the Boundary Conditions by Relaxation*

(1) *Boundaries  $x=0$ ,  $b$ .*—Here as in the previous problem, since  $u$  and  $v$  are fixed on the ends, no image values are involved and the adjacent points can be relaxed without any account being taken of residuals on the ends.

(2) *Boundary  $x+y=3b/2$ .*—The boundary conditions (iii) give:

- (a) (strain normal to the boundary)  $+\sigma$  (strain parallel to the boundary)  $=0$
- (b) (shear strain of the boundary)  $=0$ .

The displacements resolved in the  $n$  and  $s$ -directions, normal to and along the boundary as shown in Figure 10, are given by  $(u+v)/\sqrt{2}$  and  $(u-v)/\sqrt{2}$  respectively.

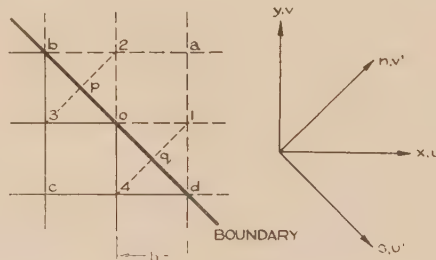


Fig. 10

The boundary conditions (a) and (b) at the point 0 may then be written in terms of finite difference approximations as follows:

$$\left. \begin{aligned} \{(u_a + v_a) - (u_c + v_c)\} + 0.3\{(u_d - v_d) - (u_b - v_b)\} &= 0 \\ \{(u_a - v_a) - (u_c - v_c)\} + \{(u_d + v_d) - (u_b + v_b)\} &= 0 \end{aligned} \right\} \dots\dots\dots (8)$$

Solving these two equations for  $u_a$  and  $v_a$  gives

$$\left. \begin{aligned} u_a &= u_c + 0.65(u_b - u_d) + 0.35(v_b - v_d) \\ v_a &= v_c - 0.65(v_b - v_d) - 0.35(u_b - u_d) \end{aligned} \right\} \dots\dots\dots (9)$$

Similarly, applying the boundary condition (iii) to points  $p$ ,  $q$  gives

$$\left. \begin{aligned} u_2 &= u_3 + 0.65(u_b - u_0) + 0.35(v_b - v_0) \\ v_2 &= v_3 - 0.65(v_b - v_0) - 0.35(u_b - u_0) \end{aligned} \right\} \dots\dots\dots (10)$$

$$\left. \begin{aligned} u_1 &= u_4 + 0.65(u_0 - u_d) + 0.35(v_0 - v_d) \\ v_1 &= v_4 - 0.65(v_0 - v_d) - 0.35(u_0 - u_d) \end{aligned} \right\} \dots\dots\dots (11)$$



For convenience the mesh-points on the boundary will be termed "primary boundary points", mesh-points in the plate, at a distance of  $h/\sqrt{2}$  from the boundary, "secondary boundary points", and  $\sqrt{2}h$  from the boundary "tertiary boundary points". Thus, referring to Figure 10,  $b, 0, d$ , are primary boundary points,  $3, 4$ , are secondary boundary points, and  $c$  is a tertiary boundary point.

Residuals at primary boundary points: After substitution of the image-values given by equations (9), (10), (11) at points 1,  $a, 2$ , the expressions for the residuals at 0 (Fig. 10) become

$$\left. \begin{aligned} F_u &= -18.22u_0 + 10.8(u_3 + u_4) + 1.365u_b - 4.745u_d \\ &\quad + 1.820v_0 + 2.60v_c - 1.165v_b - 3.255v_d \\ F_v &= -18.22v_0 + 10.8(v_3 + v_4) - 4.745v_b + 1.365v_d \\ &\quad + 1.820u_0 + 2.60u_c - 3.255u_b - 1.165u_d \end{aligned} \right\} \dots\dots\dots (12)$$

Residuals at secondary boundary points: Using the expressions given by equations (10), the image values at  $a$  (Fig. 11) are

$$\left. \begin{aligned} u_a &= u_0 + 0.65(u_2 - u_1) + 0.35(v_2 - v_1) \\ v_a &= v_0 - 0.65(v_2 - v_1) - 0.35(u_2 - u_1) \end{aligned} \right\} \dots\dots\dots (13)$$

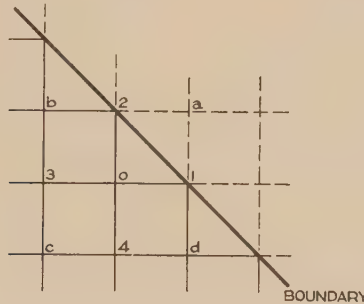


Fig. 11

Substituting these values for  $u_a, v_a$  in equations (5) the residuals at 0 (Fig. 11) become

$$\left. \begin{aligned} F_u &= -21.6u_0 + 8.455u_1 + 8u_3 + 2.345u_2 + 2.80u_4 \\ &\quad + 0.845(v_1 - v_2) + 1.30(v_0 - v_b + v_c - v_d) \\ F_v &= -21.6v_0 + 2.345v_1 + 2.80v_3 + 8.455v_2 + 8.0v_4 \\ &\quad - 0.845(u_1 - u_2) + 1.30(u_0 - u_b + u_c - u_d) \end{aligned} \right\} \dots\dots\dots (14)$$

Residuals at tertiary boundary points: These do not involve image values and are given by equations (5).

The relaxation patterns for use at primary, secondary, and tertiary boundary points obtained from equations (5), (12), (14) are given in Figures 12, 13, 14. The patterns for all other points are as shown in Figure 2.

(3) *Centre-line  $y=0$ .*—When  $u$  and  $v$  are changed at points adjacent to the centre-line the appropriate changes must be made at the image points. By symmetry  $u$  has an equal positive image in the centre-line and  $v$  an equal negative image. There is no need for special patterns as the process can best be done mentally.

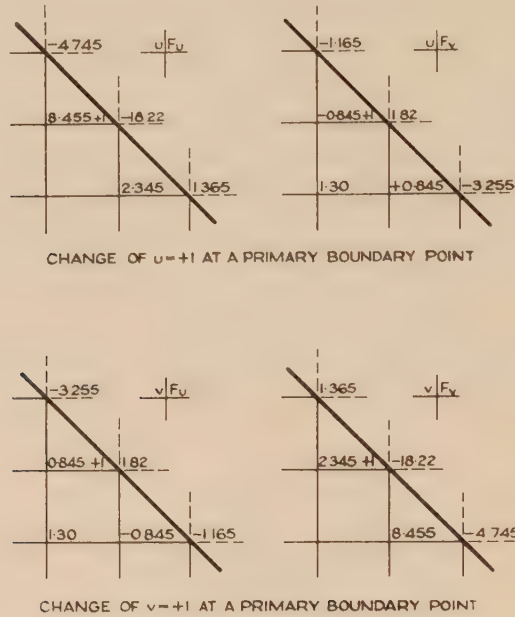


Fig. 12

### (c) Description of the Process

The process used to obtain a solution of this problem differed only in detail from that used in the first problem. The end deflection  $u_0$  was taken to be  $+16$ . For a starting assumption it was assumed that  $v$  was everywhere zero, and that the stress  $\hat{x}$  was a function of  $x$  only. Since the total load in the  $x$ -direction is a constant for any cross section of the plate perpendicular to the direction of the applied force,  $u$  is given by

$$\frac{\partial u}{\partial x}(3b-2x) = \text{constant}.$$

Solving this equation for  $u$  and applying the boundary conditions

- (i) at  $x=0$ ,  $u=0$ ,
- (ii) at  $x=b$ ,  $u=u_0=16$ ,

gives

$$u = \frac{16}{\log_e 3} \cdot \log_e \frac{3b}{3b-2x} \text{ for the starting assumption.}$$

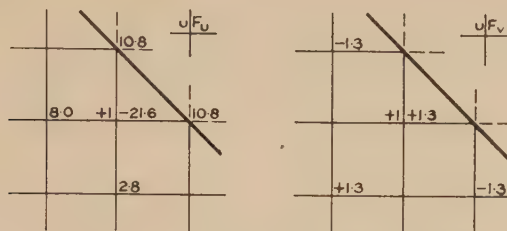
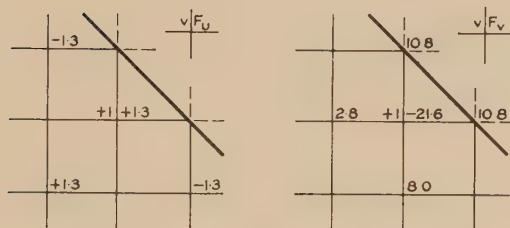
CHANGE OF  $u = +1$  AT A SECONDARY BOUNDARY POINTCHANGE OF  $v = +1$  AT A SECONDARY BOUNDARY POINT

Fig. 13

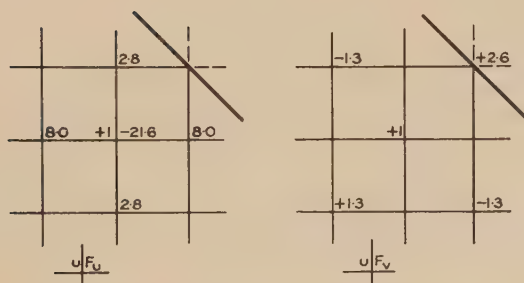
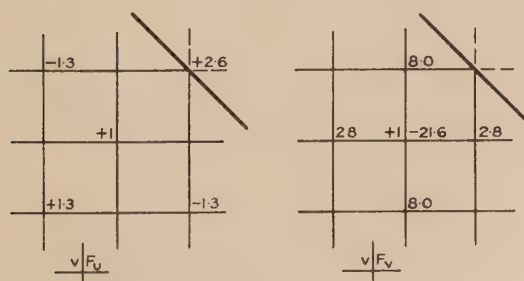
CHANGE OF  $u = +1$  AT A TERTIARY BOUNDARY POINTCHANGE OF  $v = +1$  AT A TERTIARY BOUNDARY POINT

Fig. 14

With this distribution of  $u$  and  $v$ , and using a mesh-length of  $h=b/8$ , residuals were calculated and liquidated as before. A mesh-length of  $h=b/16$  was then used and the residuals in this case reduced to such a value that a change of  $\pm 0.01$  in either  $u$  or  $v$ , in general, caused an increase in the residuals. Relaxation was not performed on a finer mesh as it was considered that the  $h=b/16$  solution was sufficiently accurate over most of the plate. However, it is realized that this solution is inaccurate in the region of the corners  $(b, \pm b/2)$  due to the stress concentrations there, and a finer mesh would be required to obtain a reasonably accurate solution near these corners. As in the first problem, it would be sufficient to consider only a localized region when using the finer mesh.

#### (d) Derivation of Stresses

The stresses  $\widehat{xx}$ ,  $\widehat{yy}$  and  $\widehat{xy}$  were calculated at all mesh-points within the boundaries of the plate using equations (1). As in the first problem, the stresses are presented in the form

$$(1-\sigma^2) \cdot \frac{b}{u_0} \cdot \frac{\text{stress}}{E}$$

The derivatives of  $u$  and  $v$  which occur in equations (1) were calculated from the finite difference approximations of equations (3), using the results obtained on the  $h=b/16$  mesh.

Since this method of calculating stresses involves image values when applied to points on the boundaries, special consideration must be given to these cases.

(1) *Boundaries*  $x=0, b$ .—The image values in the ends  $x=0, b$  were obtained from the values of  $u, v$  in the plate using a parabolic extrapolation formula. Using these image values, stresses at points on the ends  $x=0, b$  were derived. It was found that these stresses differed only slightly from the values obtained by direct extrapolation of the stress values inside the boundary.

It would, of course, be possible to use a method similar to that outlined in the first problem for the determination of the image values of  $u$  and  $v$  in the ends. However, since stress gradients are small, except in the regions of the corners  $(b, \pm b/2)$ , it was thought that extrapolation would give sufficiently accurate results.

At the corners  $(b, \pm b/2)$  there are theoretically infinite stress concentrations with consequent large stress gradients. Finite difference approximations then break down and so the stress values have been omitted in these regions.

(2) *Boundary*  $x+y=3b/2$ .—Along the edge  $x+y=3b/2$  the normal stress  $\widehat{nn}$  and the shear stress  $\widehat{ns}$  are zero. Hence for equilibrium  $\widehat{xx}=\widehat{yy}=-\widehat{xy}=\frac{1}{2} \widehat{ss}$ .

Since  $\widehat{nn}=0$ , referring to Figure 10, we have

$$\begin{aligned} \frac{\partial v^1}{\partial n} + \sigma \frac{\partial u^1}{\partial s} &= 0 \\ \therefore \widehat{ss} &= \frac{E}{1-\sigma^2} \left( \frac{\partial u^1}{\partial s} + \sigma \frac{\partial v^1}{\partial n} \right) \\ &= E \frac{\partial u^1}{\partial s} \end{aligned}$$



Now  $u' = \frac{1}{\sqrt{2}}(u-v)$ , so that the value of  $\widehat{ss}$  at the point 0 (using finite difference approximations) becomes

$$\widehat{ss} = \frac{E}{4h}(u_d - v_d - u_b + v_b)$$

$$\therefore \widehat{xx} = \widehat{yy} = -\widehat{xy} = \frac{\widehat{ss}}{2} = \frac{E}{8h}(u_d - v_d - u_b + v_b) \quad \dots\dots\dots (15)$$

This equation was used to calculate the edge stresses.

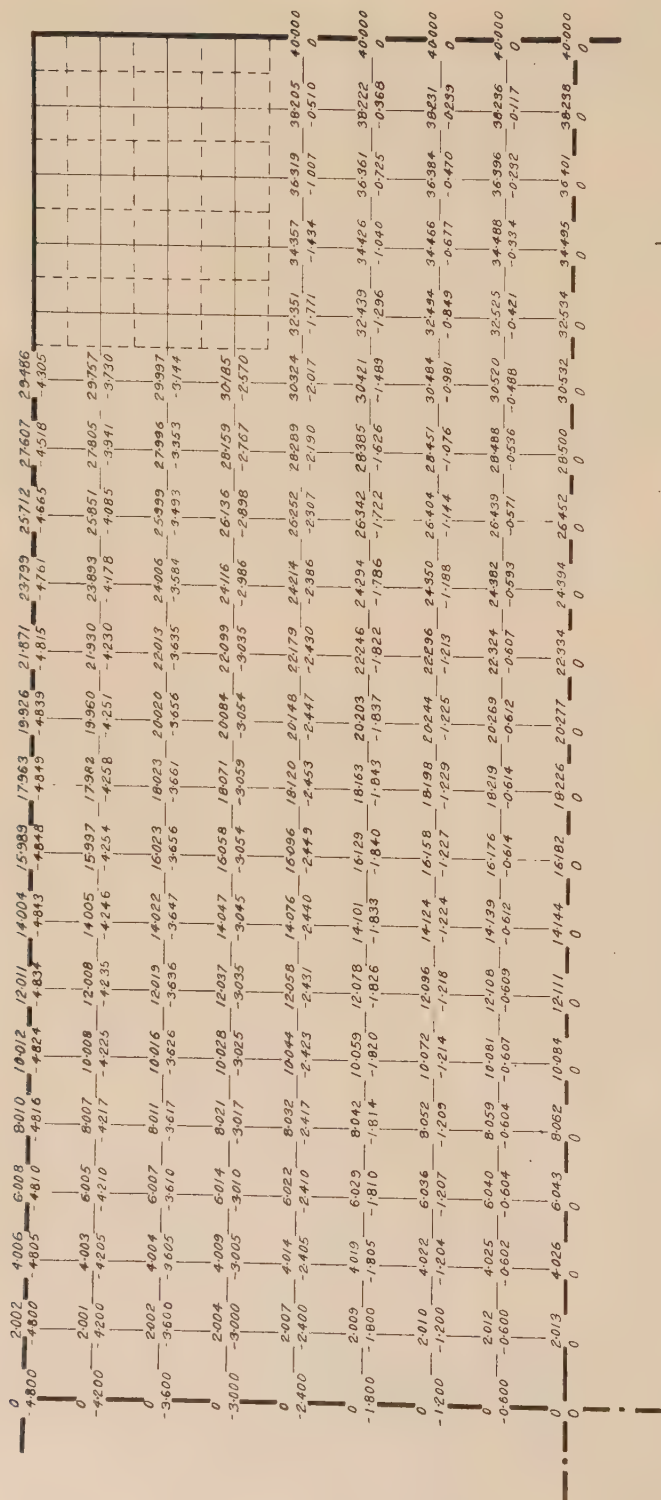
An alternative method would be to calculate the image values of  $u$  and  $v$  in the edge (using equations (10) or (11)), and thence determine the strains and the consequent stresses. When this method was used, the values of  $\widehat{xx}$ ,  $\widehat{yy}$  and  $-\widehat{xy}$  so obtained were slightly different from each other and from the common value obtained from equation (15). The discrepancy is due to the slight errors in the finite difference approximations. If relaxation had been carried out on a finer mesh, the stress values obtained by the two different methods should have converged to the same value.

## V. DISCUSSION OF RESULTS

### (a) Rectangular Plate with Free Edges and Completely Clamped Ends under Tension

The results for this problem are presented in Figures 15-17. Figures 15 and 15A give the values of the displacements  $u$  and  $v$  at the various mesh-points, for an end displacement  $u_0=40$ . Figures 16 and 16A show the values of the three stress components as derived from the displacements, and in Figure 17 the contours of the predominant stress  $\widehat{xx}$  are plotted. As mentioned previously the solution is not accurate in the region of the stress concentration and for that reason the contours have not been continued into the extreme corner. However, over the remainder of the plate the solution should be reasonably accurate on account of the relatively small stress gradients.

It will be seen from Figure 17 that the effect of clamping the ends has almost completely died away at a distance from the ends equal to the width of the plate, the variation of  $\widehat{xx}$  over this section being of the order of 1 per cent. The value of  $\widehat{xx}$  along the edge of the plate falls to a minimum from the corner to a point at a distance of approximately one-quarter of the width of the plate from the corner and then rises again to the value at the centre section. Along the centre-line parallel to the load the stress  $\widehat{xx}$  rises to a maximum at a distance of approximately half the width of the plate from the end and then falls slightly to its value at the centre section. Along this centre-line the maximum variation of  $\widehat{xx}$  is about 8.5 per cent. of its value at the centre section.

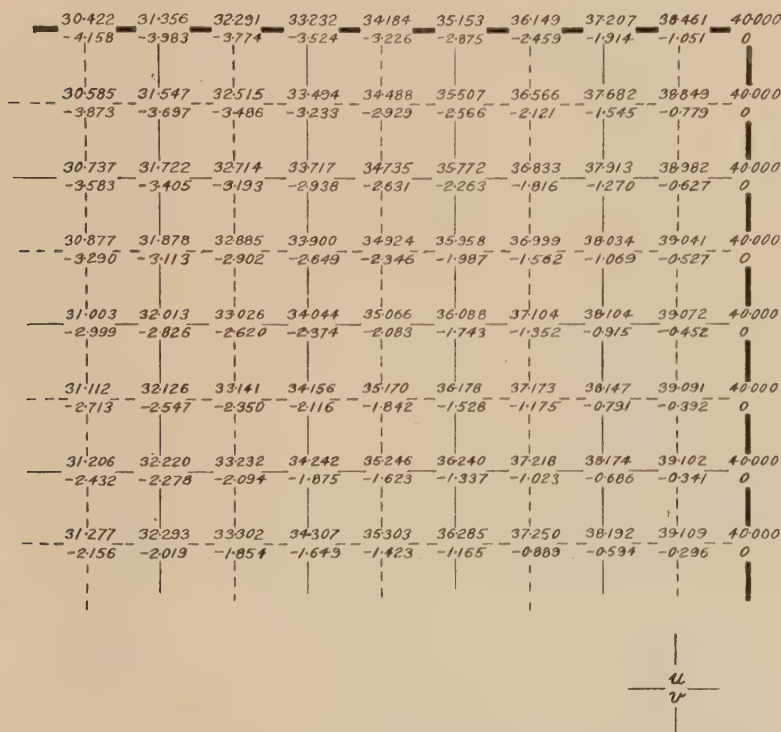


RECTANGULAR PLATE WITH CLAMPED ENDS UNDER TENSION  
VALUES OF DISPLACEMENTS  $u$  AND  $v$  WHEN  $u_0 = 40$ .  
(FOR VALUES IN THE CORNER SEE THE FINER MESH ON FIG. 15a.)

Fig. 15

(b)  $45^\circ$  Tapered Plate with Free Edges and Completely Clamped Ends under Tension

For this problem the values of the displacements  $u$  and  $v$  at the various mesh-points are given in Figure 18 for an end displacement  $u_0 = 16$ . In Figure 19 the values of the stress components are given and in Figure 20 the contours of the predominant stress  $\hat{x}$  are again plotted. As in the first problem, due to the stress concentration at the corner, the contours are discontinued as they approach this region. However, over the remaining region the solution should again be accurate.



RECTANGULAR PLATE WITH CLAMPED ENDS UNDER TENSION. VALUES OF DISPLACEMENTS  $u$  AND  $v$  IN THE REGION OF THE CORNER. (SUPPLEMENTARY TO FIG. 15.)

Fig. 15a

The main point of interest is the fact that along the centre-line the stress  $\hat{x}$  remains remarkably constant. At first sight it would be expected that the stress would rise steadily from the wider end to the narrower end with the extreme stresses in the ratio of approximately 3/1.



RECTANGULAR PLATE WITH CLAMPED ENDS UNDER TENSION  
VALUES OF STRESS COMPONENTS  $\bar{x}\bar{x}$ ,  $\bar{y}\bar{y}$  AND  $\bar{x}\bar{y}$   
EXPRESSED IN THE NON-DIMENSIONAL FORM

$$(1-\sigma^2)^{\frac{1}{2}} \frac{p}{E}$$

(FOR VALUES IN THE CORNER SEE THE FINER MESH ON FIG. 16a.)

Fig. 16



## VI. CONCLUSIONS

In the two problems solved here the mixed boundary conditions present no particular difficulty. The method has not been extended to deal with the case of curved boundaries but there seems to be no fundamental reason why this should not be possible.

-.6804	-.6828	-.6892	-.6992	-.7152	-.7476	-.8416	-.10168	-.0000
-.0000	-.0000	-.0000	-.0000	-.0000	-.0000	-.0000	-.0000	-.0000
-.7028	-.7092	-.7188	-.7336	-.7576	-.7928	-.8360	-.8764	-.9852
-.0004	-.0012	-.0024	-.0036	-.0044	-.0040	-.0164	-.1084	-.2956
-.0028	-.0056	-.0100	-.0164	-.0264	-.0472	-.0892	-.1432	-.3320
-.7208	-.7280	-.7384	-.7520	-.7696	-.7892	-.8024	-.8044	-.7264
-.0036	-.0060	-.0088	-.0136	-.0200	-.0332	-.0676	-.1496	-.2180
-.0084	-.0136	-.0220	-.0336	-.0508	-.0784	-.1172	-.1508	-.1388
-.7336	-.7400	-.7480	-.7576	-.7676	-.7748	-.7740	-.7652	-.7584
-.0092	-.0136	-.0192	-.0276	-.0412	-.0636	-.1032	-.1660	-.2276
-.0136	-.0212	-.0320	-.0464	-.0656	-.0904	-.1180	-.1372	-.1472
-.7412	-.7460	-.7520	-.7572	-.7600	-.7600	-.7540	-.7420	-.7236
-.0164	-.0228	-.0316	-.0436	-.0608	-.0872	-.1248	-.1736	-.2172
-.0136	-.0276	-.0392	-.0540	-.0716	-.0916	-.1100	-.1212	-.1196
-.7460	-.7488	-.7516	-.7536	-.7524	-.7480	-.7396	-.7280	-.7148
-.0244	-.0332	-.0440	-.0584	-.0780	-.1048	-.1384	-.1760	-.2144
-.0220	-.0316	-.0436	-.0572	-.0720	-.0872	-.1000	-.1064	-.1068
-.7472	-.7492	-.7496	-.7488	-.7452	-.7392	-.7300	-.7188	-.7064
-.0320	-.0444	-.0548	-.0720	-.0916	-.1176	-.1472	-.1808	-.2120
-.0240	-.0340	-.0448	-.0568	-.0692	-.0804	-.0892	-.0936	-.0932

RECTANGULAR PLATE WITH CLAMPED ENDS UNDER TENSION. VALUES OF STRESS COMPONENTS  $\bar{x}\bar{x}$ ,  $\bar{y}\bar{y}$  AND  $\bar{x}\bar{y}$  IN THE REGION OF THE CORNER, EXPRESSED IN THE NON-DIMENSIONAL FORM

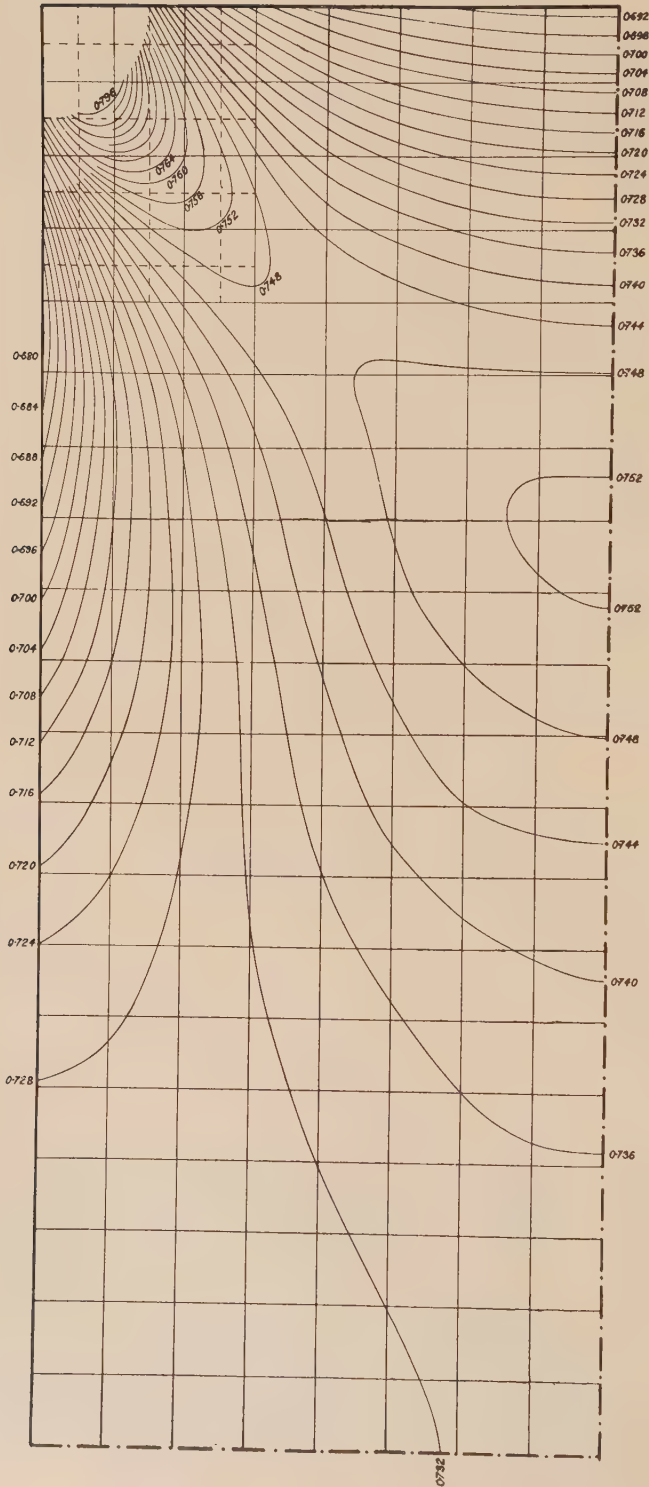
$$(1-\sigma^2) \frac{b}{a} \frac{\text{STRESS}}{E}$$

(SUPPLEMENTARY TO FIG. 16)



Fig. 16A

In view of the fact that the relaxation process gives a much more rapid convergence for the two partial differential equations involving displacements than for the equivalent equation involving the stress function, it is suggested that, in spite of the more complicated boundary conditions, it may be an advantage to work in terms of displacements for all two-dimensional stress problems. The extra work introduced in keeping account of two residuals in place of only the one when a stress function solution is used is more than counteracted by increased convergence. Also the displacements have a more readily appreciated physical significance than a stress function, a fact which is of considerable assistance in formulating the first guess at the distribution.



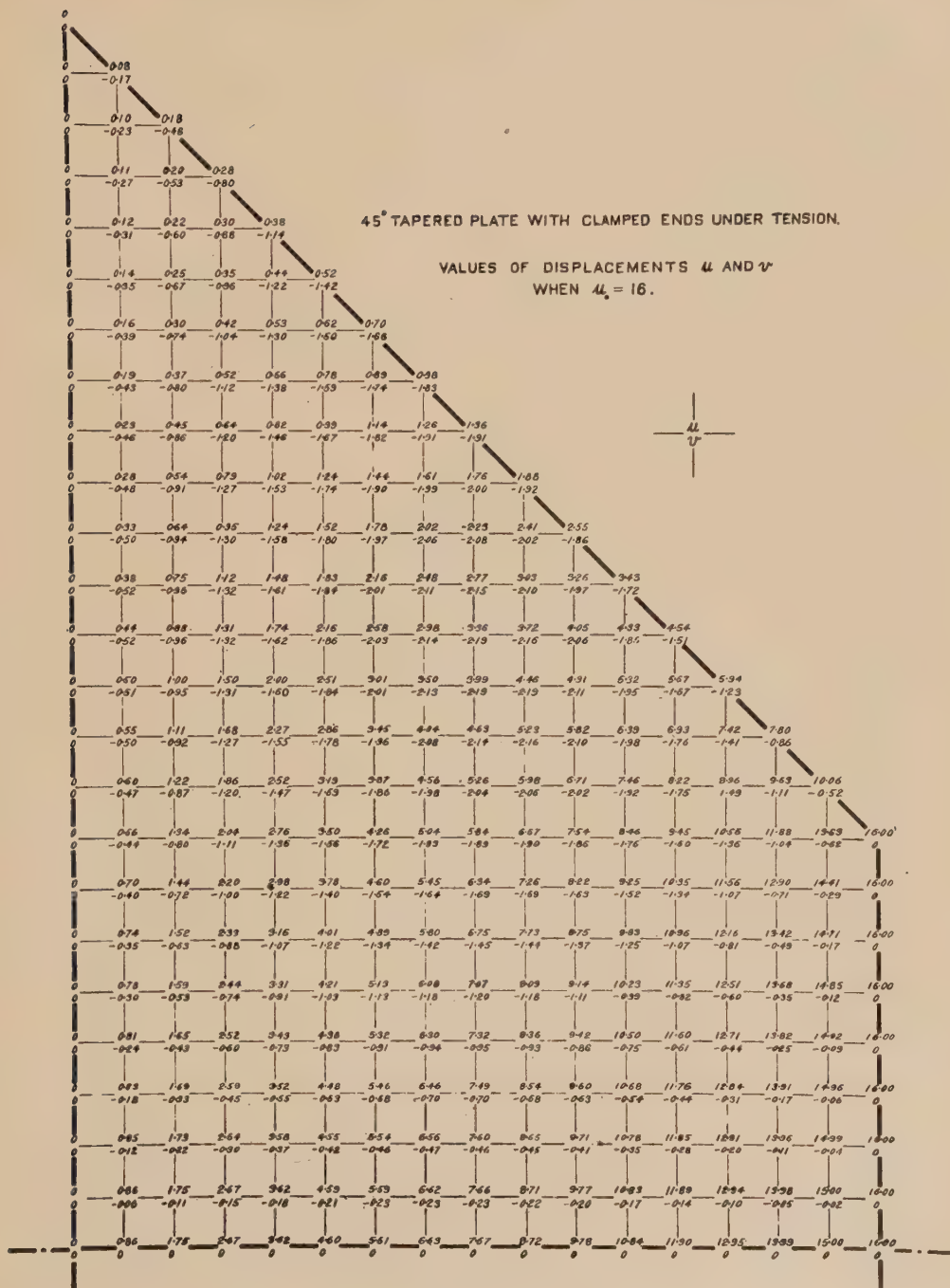


Fig. 18

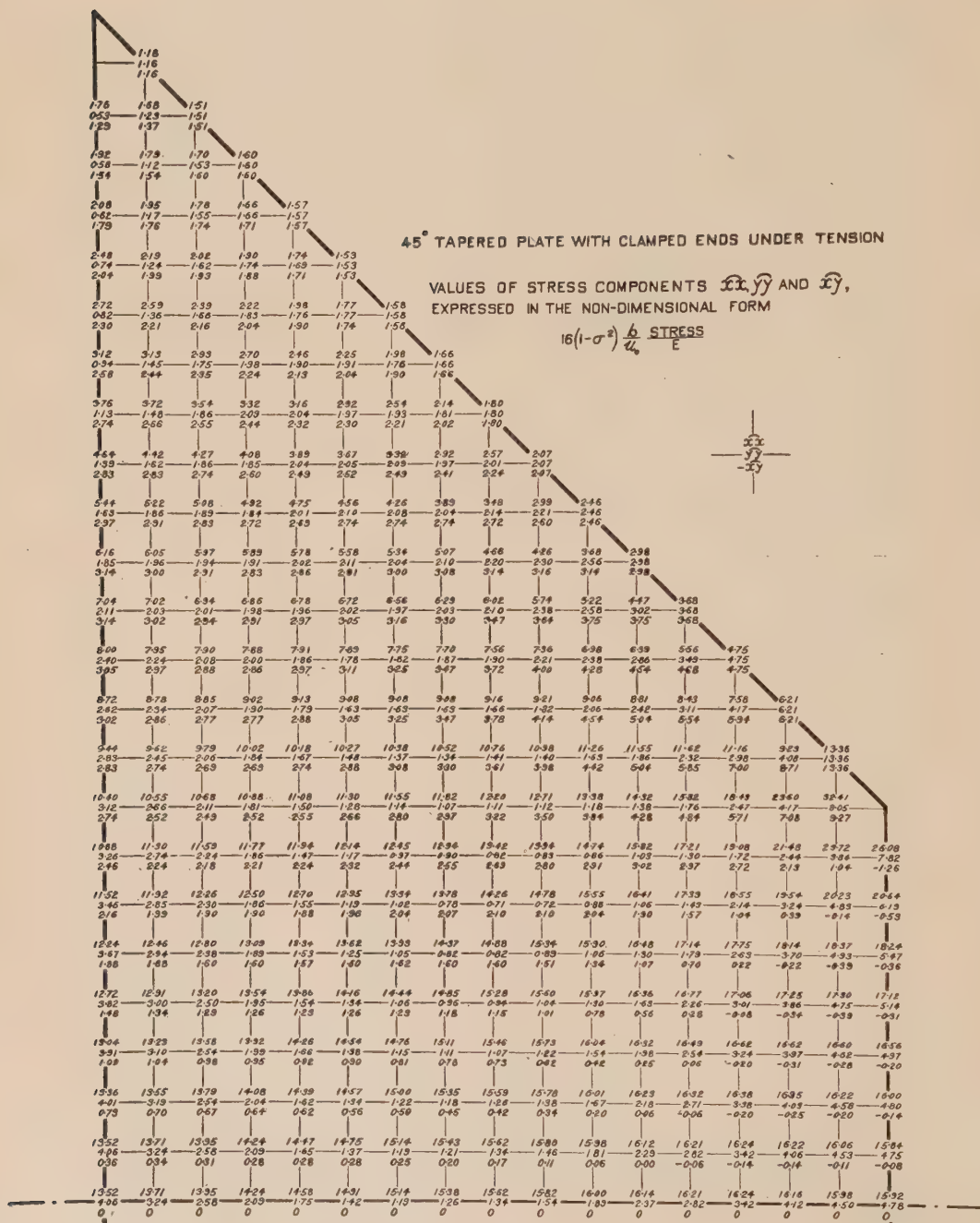


Fig. 19



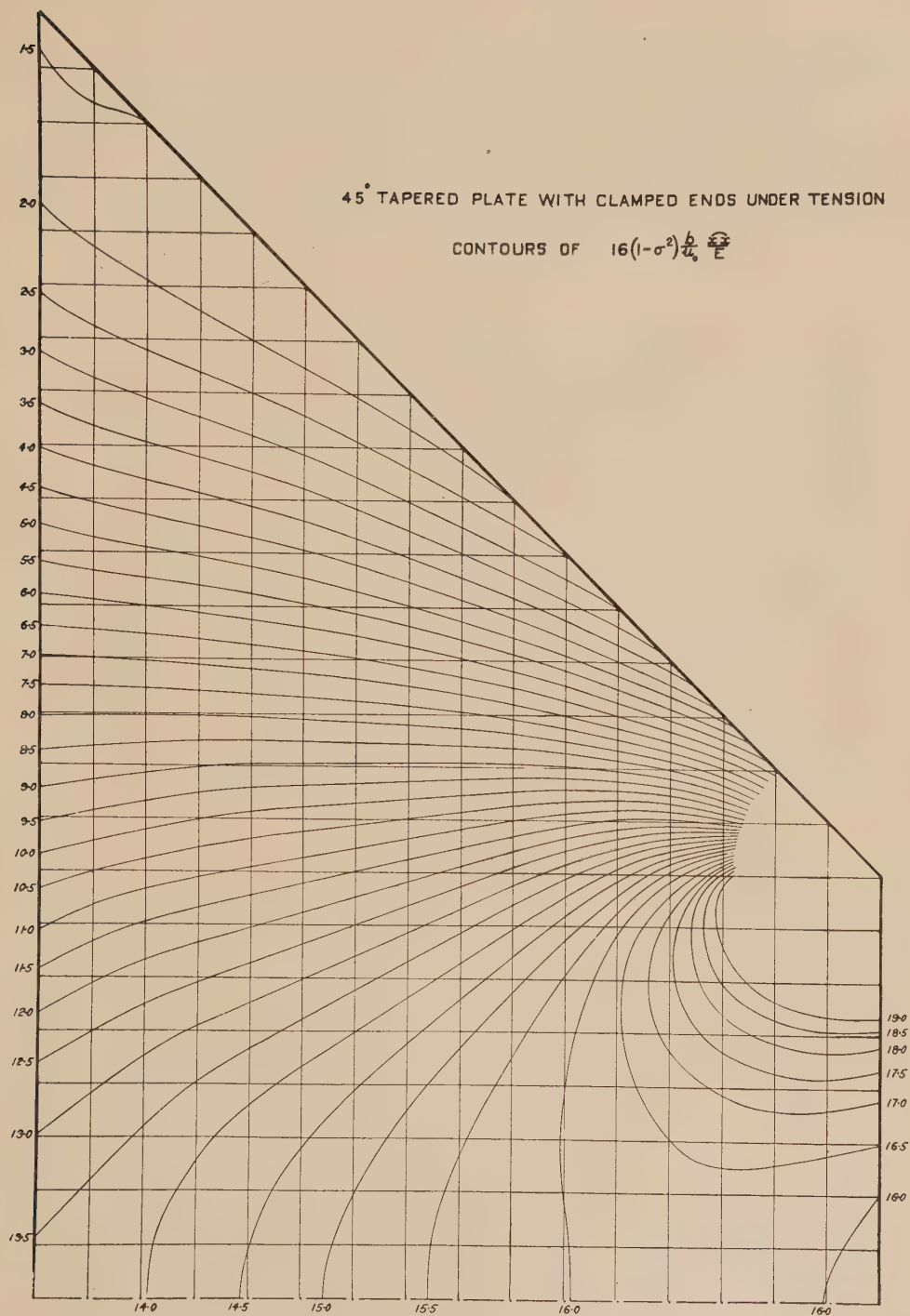


Fig. 20

## VII. REFERENCES

- (1) SOUTHWELL, R. V.—“Relaxation Methods in Theoretical Physics.” (Oxford University Press, 1946.)
- (2) CHRISTOPHERSON, D. G., and SOUTHWELL, R. V.—Relaxation methods applied to engineering problems. III. Problems involving two independent variables. *Proc. Roy. Soc. A* **168** : 317-50 (1938).
- (3) FOX, L., and SOUTHWELL, R. V.—Id. VIIA. Biharmonic analysis as applied to the flexure and extension of flat elastic plates. *Philos. Trans. A* **239** : 419-60 (1945).

# LOCAL ISOTROPY IN THE TURBULENT WAKE OF A CYLINDER

By A. A. TOWNSEND\*

[Manuscript received April 29, 1948]

## Summary

To investigate the validity of Kolmogoroff's theory of local isotropy, a series of turbulence measurements have been made in the wake of a circular cylinder, and the results have been compared with the predictions of the theory. Using a cylinder of 0.953 cm. diameter in an air-stream of velocity 1,280 cm. sec.<sup>-1</sup>, measurements have been made of the mean squares of the spatial derivatives in the mean-stream direction of the three components of the turbulent velocity fluctuation, and also of the skewness and flattening factors of the statistical distributions of these derivatives. Observations were taken at three traverses across the wake, respectively at 80, 120, and 160 cylinder diameters down-stream from the cylinder. Except in the immediate neighbourhood of the wake centre, the turbulent flow is observed to be intermittent, consisting of regions of fully developed turbulent flow separated by comparatively sharp boundaries from regions of almost completely laminar motion. If an "intermittency factor" is introduced to describe this phenomenon, and if inside each turbulent region local isotropy exists, then all the experimental results are consistent with the theory of local isotropy, and in agreement with previous measurements in flows possessing ordinary isotropy. It is concluded that, within the boundaries of the turbulent regions, local isotropy in the sense used by Kolmogoroff exists, and that the theory is applicable to this example of shear flow. The general applicability of the theory to turbulent shear flow at high Reynolds numbers must be considered very probable.

## I. INTRODUCTION

A quantitative description of the small-scale structure of fully developed turbulent flow has been provided by the theory of local isotropy, originated by Kolmogoroff(1, 2), and based on the familiar notion of the instability of large eddies. As originally presented by Kolmogoroff, experimental evidence for the truth of the theoretical predictions was slight and not entirely conclusive, but improved and extended measurements at Cambridge have led to a fairly definite verification of the main predictions of the theory in the simple case of turbulent flow which also possesses ordinary isotropy(3). In its original form the theory of local isotropy is equally applicable to shear flow, provided the Reynolds number of the mean flow is sufficiently large, and it is naturally of great interest that this prediction should be checked. Previous measurements by the author(4) of turbulent intensities in the wake of a cylinder showed that the total energy dissipation due to viscosity is in rough agreement with the predictions of the theory, but the comparison is probably not very sensitive to

\* Emmanuel College, Cambridge.

deviations from local isotropy. To extend these results, a general investigation into the detailed structure of the turbulent wake behind a circular cylinder has been undertaken, and a number of quantities have been measured whose relative magnitudes are predicted by the theory of local isotropy. These results are used below to verify the existence of local isotropy in this particular example of shear flow with free turbulence.

## II. PREDICTIONS OF THE THEORY OF LOCAL ISOTROPY

It is well established that the high energy dissipation characteristic of turbulent flow is a consequence of the instability of the larger eddies, which lose energy by breaking up into new eddies an order of magnitude smaller, which are themselves unstable. Still smaller eddies are produced in this way, until eddies are produced so small that viscosity is the controlling factor in their motion and energy is only lost by viscous dissipation. Kolmogoroff has considered the whole sequence of eddies as being in moving equilibrium, the energy associated with a certain range of eddy sizes remaining nearly constant. In such a sequence, most of the turbulent energy resides in the largest eddies whose size is determined by the characteristics of the mean flow. Furthermore, transfer of energy from mean flow to eddy motion only takes place efficiently for eddies of size comparable with the dimensions of the mean flow. So reasonably small eddies have no direct connection with the mean flow, being derived from the mean flow through a long chain of transitions, and it is expected that eddies small compared with the dimensions of the mean flow will be statistically isotropic. It is of course necessary that the Reynolds number should be high in order that eddies can exist small compared with the dimensions of the mean flow.

To express these notions quantitatively, Kolmogoroff uses the velocity difference  $(u_2 - u_1)$ , where  $u_2, u_1$  are parallel velocity components at points  $P_1$  and  $P_2$ , distant  $r$  apart on a line parallel to the direction of the components, as a measure of the velocity associated with eddies of size  $r$ . Eddies much larger than  $r$  will contribute little to  $(u_2 - u_1)$  because their velocity gradient is small, and eddies much smaller than  $r$  similarly cannot have much influence on  $(u_2 - u_1)$  because of their small extent. Then  $B_{dd} = \overline{(u_2 - u_1)^2}$  may be termed the intensity of turbulence associated with eddies of scale  $r$ , and  $(u_2 - u_1)$  is isotropic in the sense that the mean values of functions of  $(u_2 - u_1)$  are invariant under rotations and reflections of the axes of reference, if  $r$  is sufficiently small. In particular, defining  $B_{nn} = \overline{(v_2 - v_1)^2}$ , where  $v_1, v_2$  are parallel velocity components perpendicular to  $P_1P_2$ , application of the continuity condition for incompressible flow leads to the relation

$$B_{nn} = B_{dd} + \frac{1}{2}r \frac{\partial B_{dd}}{\partial r} \dots\dots\dots (1)$$

Then, allowing  $r$  to approach zero,

$$\left(\frac{\partial u}{\partial r}\right)^2 = \frac{1}{2} \left(\frac{\partial v}{\partial r}\right)^2 \dots\dots\dots (2)$$



and the vorticity fluctuations are statistically isotropic. Other consequences are that

$$S_0(u) = \left( \overline{\frac{\partial u}{\partial r}} \right)^3 / \left[ \left( \overline{\frac{\partial u}{\partial r}} \right)^2 \right]^{3/2}$$

$$T_0(u) = \left( \overline{\frac{\partial u}{\partial r}} \right)^4 / \left[ \left( \overline{\frac{\partial u}{\partial r}} \right)^2 \right]^2$$

and

$$T_0(v) = \left( \overline{\frac{\partial v}{\partial r}} \right)^4 / \left[ \left( \overline{\frac{\partial v}{\partial r}} \right)^2 \right]^2$$

should be all absolute constants, independent of the particular flow. Also due to the local isotropy

$$S_0(v) = \left( \overline{\frac{\partial v}{\partial r}} \right)^3 / \left[ \left( \overline{\frac{\partial v}{\partial r}} \right)^2 \right]^{3/2} = 0.$$

A number of other relations may be deduced, for which reference should be made to a paper by Batchelor(5), who treats the theory in more detail than is possible here.

Confirmation of the validity of the theory has been obtained so far only when ordinary isotropy already exists in the flow, and, in this special case, the invariance of  $S_0(u)$  and  $T_0(u)$  has been confirmed. From the mean of many measurements of these quantities, it is found that

$$\begin{aligned} S_0(u) &= -0.38 \\ T_0(u) &= 3.5 \end{aligned}$$

and it is expected that similar results would be obtained in non-uniform flow if local isotropy exists.

### III. EXPERIMENTAL DETAILS

The measurements were made in the wind-tunnel at the Cavendish Laboratory, Cambridge, which has been described previously(6). Briefly, it is a closed return tunnel with a contraction ratio of 9 : 1, and the root-mean-square value of the down-stream component of the free stream turbulence is approximately 0.06 per cent. The working section measures 38.1 by 38.1 cm.<sup>2</sup> in section, and 160 cm. long, and diverges slightly to obtain substantially zero velocity gradient along the section in the down-stream direction. The wake is produced behind a circular cylinder of diameter 0.953 cm., spanning the tunnel at the entrance to the working section, and measurements have been made at three distances down stream, respectively 80, 120, and 160 cylinder diameters, at a wind-speed of 1,280 cm. sec.<sup>-1</sup>. The Reynolds number formed from the cylinder diameter and the wind-speed is 8,400.

It is necessary now to define a system of axes of reference. Taking the origin  $O$  at the centre of the cylinder, then  $Oz$  is the axis of the cylinder,  $Ox$  is parallel to the mean stream, and  $Oy$  is at right angles to the plane of the wake. Turbulent velocity components will be denoted by  $u$ ,  $v$ , and  $w$ , the

mean stream velocity by  $U$ , and the undisturbed mean stream velocity by  $U_0$ . In these experiments, the distribution of the following quantities in the  $y$  direction has been measured for the three values of  $x$ .

$$\begin{array}{ccc} \overline{\left(\frac{\partial u}{\partial x}\right)^2}, & \overline{\left(\frac{\partial v}{\partial x}\right)^2}, & \overline{\left(\frac{\partial w}{\partial x}\right)^2}, \\ S_0(u), & T_0(u), & T_0(v), \quad T_0(w). \end{array}$$

The fundamental measuring instrument is the hot-wire anemometer, used both in the single-wire form to measure  $u$  (down-stream) fluctuations, and in the double  $V$ -wire form to measure  $v$  or  $w$  (cross-stream) fluctuations. The  $V$ -wire consists of two platinum Wollaston wires of 0.00025 cm. diameter, each about 1.2 mm. long, arranged in the form of a  $V$  with the vertex pointing upstream. The sensitive volume of the  $V$ -wire is greater than that of a similar  $X$ -wire, but the turbulent flow in these experiments has a sufficiently large scale to make errors due to this cause reasonably small.

The electrical fluctuations from the hot-wire bridge are amplified, compensated for the thermal capacity of the wire, and then differentiated with respect to time in an amplifier of the type previously described by the author(7). As explained then, this time differentiation is equivalent to differentiation with respect to  $x$ ,

$$\frac{\partial}{\partial x} = -\frac{1}{U} \frac{\partial}{\partial \tau}$$

where  $\tau$  is the time variable, provided that the turbulence level is small. Then

$$\begin{aligned} \overline{\left(\frac{\partial u}{\partial x}\right)^2} &= \frac{1}{U^2} \overline{\left(\frac{\partial u}{\partial \tau}\right)^2} \\ \overline{\left(\frac{\partial v}{\partial x}\right)^2} &= \frac{1}{U^2} \overline{\left(\frac{\partial v}{\partial \tau}\right)^2} \\ \overline{\left(\frac{\partial w}{\partial x}\right)^2} &= \frac{1}{U^2} \overline{\left(\frac{\partial w}{\partial \tau}\right)^2} \end{aligned}$$

and the mean squares of the time derivatives are measured in the usual way with a vacuum thermojunction and a millivoltmeter.

The remaining quantities are non-dimensional and have been measured from the differentiated amplifier output using a circuit fundamentally similar to that previously used in the measurement of triple correlations (Fig. 1). Essentially, the input is accurately balanced with respect to ground, and then an output proportional to the instantaneous square of the input signal is produced by the action of two triodes which have been adjusted so that they are working on the parabolic part of their grid-plate transfer characteristics. Since resistance-capacity coupling is used, the mean value of the output is zero, and the available output must be

$$\alpha(e^2 - \overline{e^2})$$

where  $\alpha$  is a constant

and  $e$  is the input voltage.

If the original input is added to this "squared" signal, and the mean square value of the sum measured by the ordinary method, the millivoltmeter deflection is

$$\begin{aligned}\theta_+ &= A[e + \alpha(e^2 - \bar{e}^2)]^2 \\ &= A[\bar{e}^2 + \alpha^2(\bar{e}^4 - (\bar{e}^2)^2) + 2\alpha\bar{e}^3]\end{aligned}$$

where  $A$  is a constant, characteristic of the circuit.

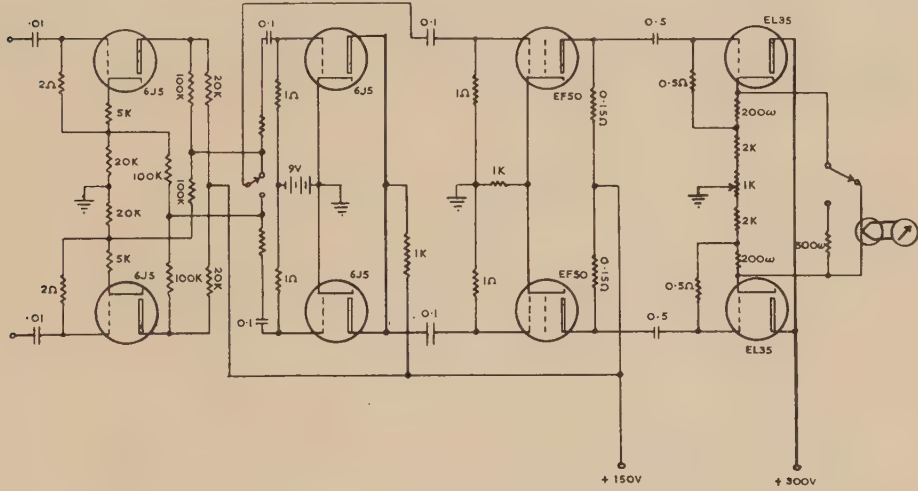


Fig. 1.—Measurement of shape factors.

If the sign of the input  $e$  is now reversed, then the millivoltmeter deflection is

$$\theta_- = A[\bar{e}^2 + \alpha^2(\bar{e}^4 - (\bar{e}^2)^2) - 2\alpha\bar{e}^3]$$

and

$$\theta_+ - \theta_- = 4A\alpha\bar{e}^3.$$

The value of  $A\alpha$  may be obtained by using a signal of known skewness factor. The standard signal is

$$e = e_0(\sqrt{2} \cos pt + \cos 2 pt)$$

for which  $\bar{e}^3 = \sqrt{\frac{2}{3}}(\bar{e}^2)^{3/2}$ . Such a signal is readily constructed electrically. An alternative method of calibration is to measure

$$\frac{1}{4} \frac{\theta_+ - \theta_-}{\sqrt{\theta_1 \theta_2}} = \frac{\bar{e}^3}{(\bar{e}^2)^{3/2} \sqrt{\frac{\bar{e}^4}{(\bar{e}^2)^2} - 1}} = \frac{\beta_3(e)}{\sqrt{\beta_4(e) - 1}}$$

where  $\theta_1 = A\bar{e}^2$ , i.e. the deflection with the original input only,

$\theta_2 = A\alpha^2(\bar{e}^4 - (\bar{e}^2)^2)$ , i.e. the deflection with the squared input only,

$$\beta_4(e) = \bar{e}^4 / (\bar{e}^2)^2$$

and  $\beta_3(e) = \bar{e}^3 / (\bar{e}^2)^{3/2}$ .

This requires a knowledge of  $\beta_4(e)$ , which is given by

$$\beta_4(e) = 1 + \frac{\theta_2}{\theta_1^2} \frac{\alpha^2}{A}$$

If the input  $e$  is made a pure sinusoidal voltage, then

$$\beta_4(e) = 1.5$$

and the factor  $\alpha^2/A$  may then be determined and applied in the computation of experimental values of  $\beta_4(e)$ . Hence the skewness factor  $\beta_3(e)$  and the flattening factor  $\beta_4(e)$  may both be computed from measurements of the four mean square intensities,  $\theta_+$ ,  $\theta_-$ ,  $\theta_1$ ,  $\theta_2$ . In the experiments,  $e$  is proportional to either  $\partial u/\partial x$ ,  $\partial v/\partial x$  or  $\partial w/\partial x$ , and then, if  $e \propto \partial u/\partial x$ ,

$$\beta_3(e) = S_0(u)$$

$$\beta_4(e) = T_0(u)$$

and so on.

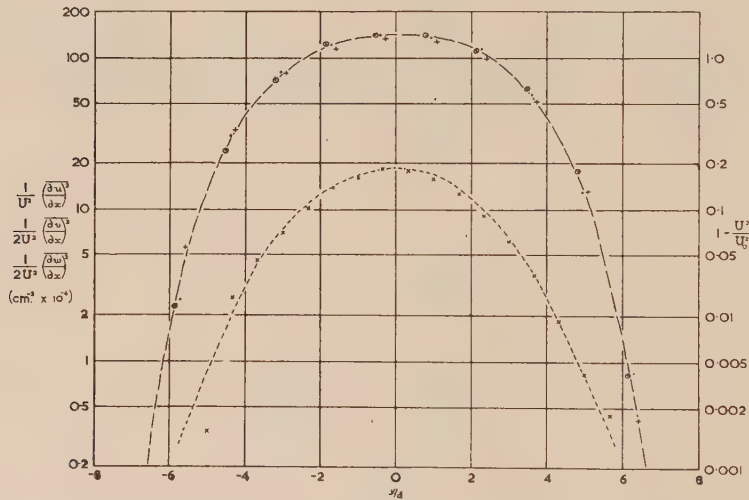


Fig. 2.—Intensities of derivatives at  $x/d=80$

•	$u$
+	$v$
⊙	$w$
×	$1 - \frac{U}{U_0^2}$

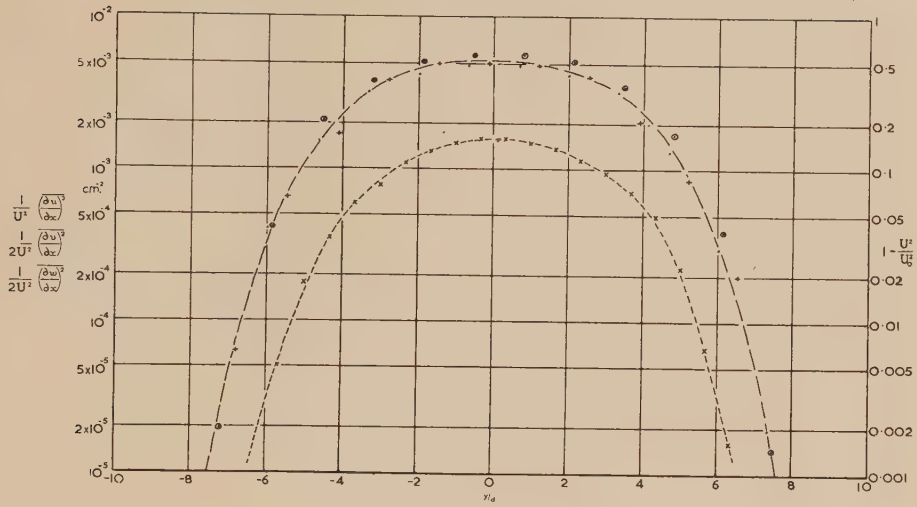
#### IV. RESULTS

In Figures 2-4, the observed values of  $\frac{1}{U^2} \overline{\left(\frac{\partial u}{\partial x}\right)^2}$ ,  $\frac{1}{2U^2} \overline{\left(\frac{\partial v}{\partial x}\right)^2}$ , and  $\frac{1}{2U^2} \overline{\left(\frac{\partial w}{\partial x}\right)^2}$  are plotted as functions of  $y/d$  ( $d$  is the cylinder diameter) for the three values of  $x/d$ . To show more clearly how the relations

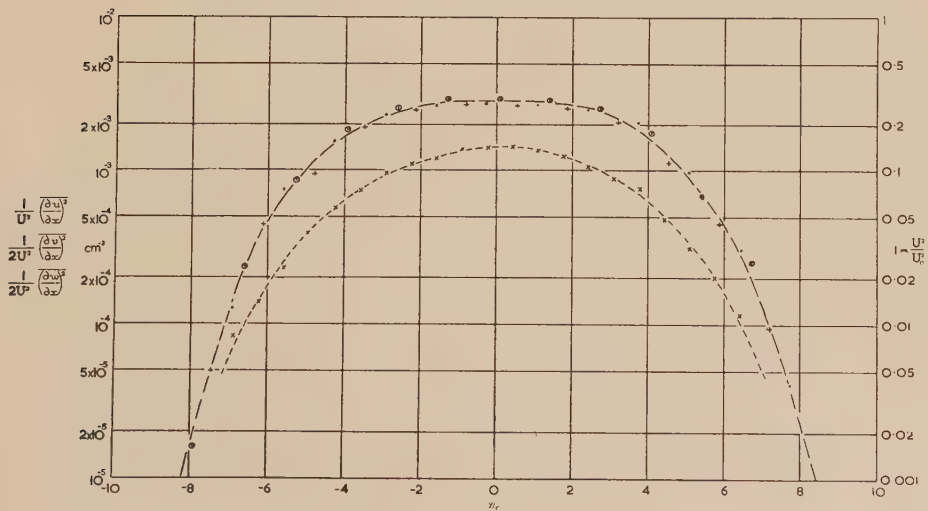
$$\overline{\left(\frac{\partial u}{\partial x}\right)^2} = \frac{1}{2} \overline{\left(\frac{\partial v}{\partial x}\right)^2} = \frac{1}{2} \overline{\left(\frac{\partial w}{\partial x}\right)^2}$$

are satisfied over a very wide range of intensity, the intensity scale is logarithmic.



Fig. 3.—Intensities of derivatives at  $x/d=120$ 

$$\left[ \begin{array}{c} \cdot \\ + \\ \odot \\ \times \end{array} \begin{array}{c} u \\ v \\ w \\ 1 - \frac{U^2}{U_0^2} \end{array} \right]$$

Fig. 4.—Intensities of derivatives at  $x/d=160$ 

$$\left[ \begin{array}{c} \cdot \\ + \\ \odot \\ \times \end{array} \begin{array}{c} u \\ v \\ w \\ 1 - \frac{U^2}{U_0^2} \end{array} \right]$$

For comparison, the mean velocity distribution is also plotted on a different scale. Except near the wake centre, the turbulent flow is definitely intermittent, with intervals of nearly laminar flow separating "bursts" of fully developed

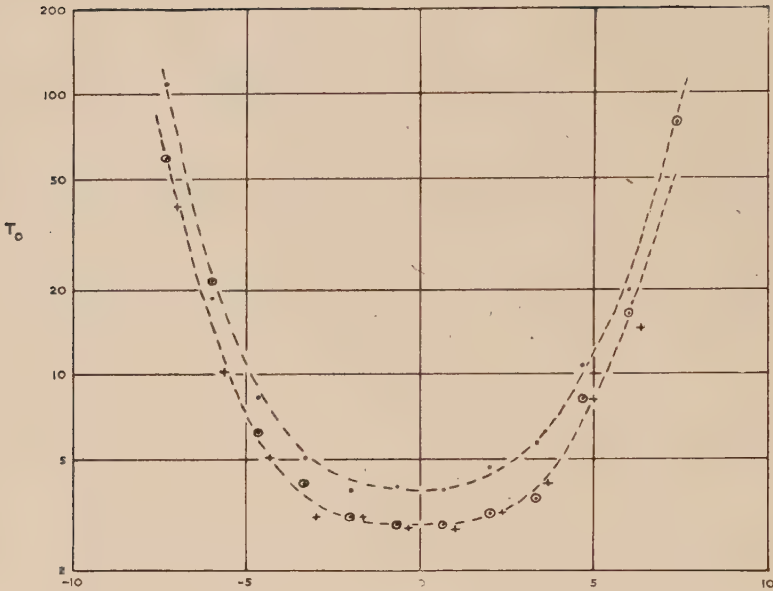


Fig. 5.—Variation of flattening factors  $\left[ \begin{array}{c} \cdot T_0(u) \\ + T_0(v) \\ \odot T_0(w) \end{array} \right]$

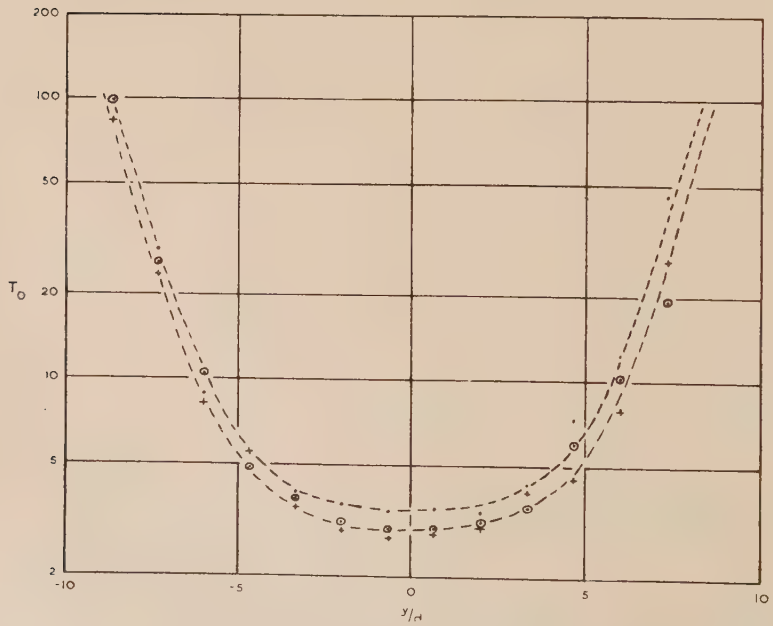


Fig. 6.—Variation of flattening factors  $\left[ \begin{array}{c} \cdot T_0(u) \\ + T_0(v) \\ \odot T_0(w) \end{array} \right]$

The predictions of the original theory of local isotropy are therefore only verified for the central part of the wake.

### V. DISCUSSION OF RESULTS

The results of these experiments may be summarized by saying that those predictions of the theory of local isotropy which concern all the quantities measured in these experiments are verified for the central part of the wake only. On the other hand, predictions concerning the relative magnitudes of mean squares of velocity derivatives are confirmed over the entire observable wake,

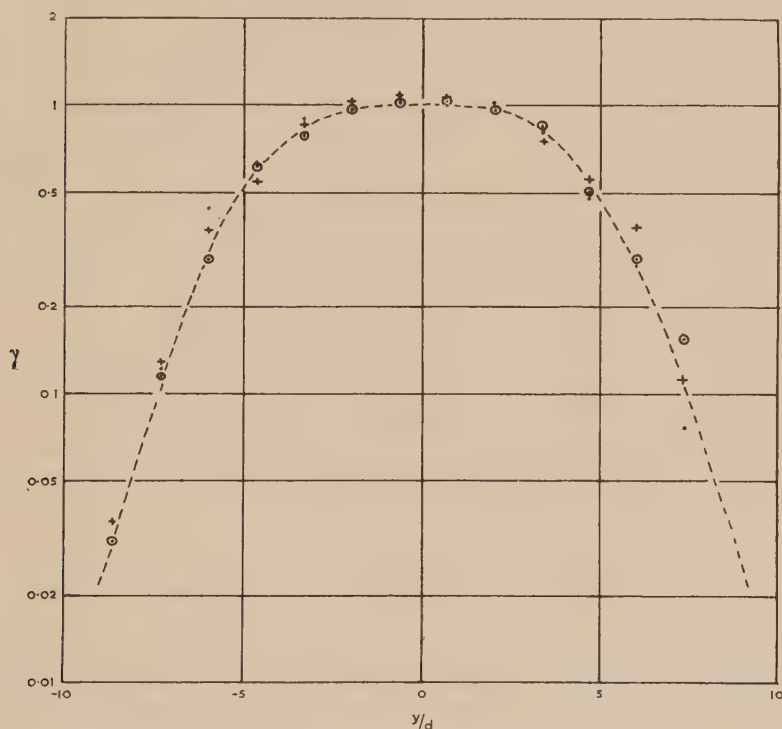


Fig. 10.—Variation of intermittency factor for  $x/d=120$   $\left[ \begin{array}{c} \cdot \quad T_0(u) \\ + \quad T_0(v) \\ \odot \quad T_0(w) \end{array} \right]$

although the predicted constancy of shape factors such as  $T_0(u)$  fails to appear. It is natural to assume that the observed intermittent nature of the turbulent flow is responsible for the failure of local isotropy in the outer portions of the wake, and it seems to be possible to explain the observed behaviour in terms of a modified local isotropy. It is observed that there is a relatively sharp division between regions of fully developed turbulence and regions of nearly laminar flow, and it may be postulated that the regions of fully developed turbulence have existed for a sufficiently long time to allow the establishment of local isotropy over substantially the whole region. The flow picture is of an intermittent

isotropy about the relative intensities of turbulent velocity derivatives are quite accurate over substantially the whole wake, including a considerable range of intensity and mean flow conditions.

The flattening factors  $T_0(u) = \left( \overline{\frac{\partial u}{\partial x}} \right)^4 / \left[ \left( \overline{\left( \frac{\partial u}{\partial x} \right)^2} \right)^2 \right]$ ,  $T_0(v) = \left( \overline{\frac{\partial v}{\partial x}} \right)^4 / \left[ \left( \overline{\left( \frac{\partial v}{\partial x} \right)^2} \right)^2 \right]$ , and  $T_0(w) = \left( \overline{\frac{\partial w}{\partial x}} \right)^4 / \left[ \left( \overline{\left( \frac{\partial w}{\partial x} \right)^2} \right)^2 \right]$ , are shown in Figures 5-7. All traverses of the wake show a similar variation, viz. nearly constant values close to the wake centre, and a rapid increase to very high values as the edge of the wake is approached. Near the wake centre, the mean value of  $T_0(u)$  is approximately 3.5, and the mean value of  $T_0(v)$  or  $T_0(w)$  is approximately 3.0.

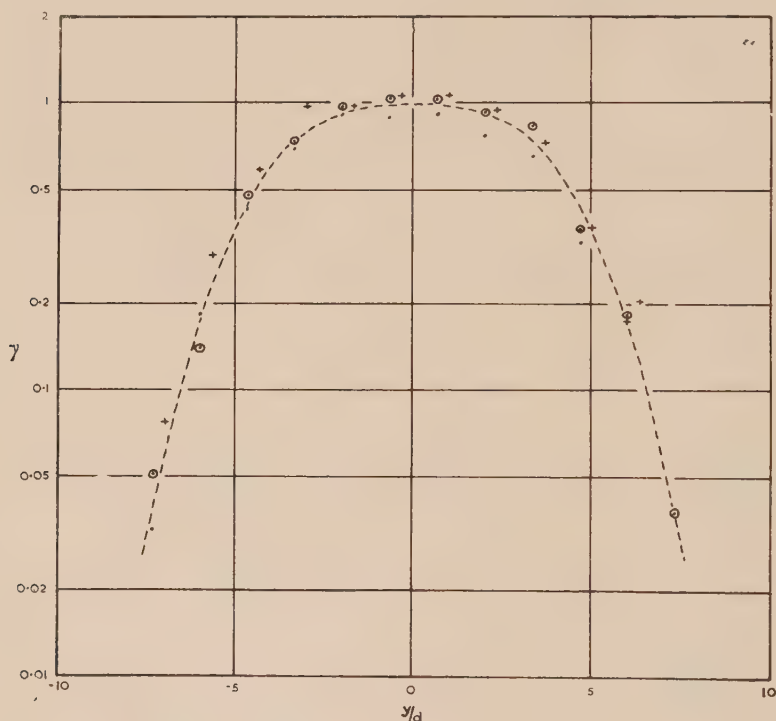


Fig. 9.—Variation of intermittency factor for  $x/d=80$

$$\begin{bmatrix} \cdot & T_0(u) \\ + & T_0(v) \\ \odot & T_0(w) \end{bmatrix}$$

The skewness factor of  $\partial u / \partial x$ ,  $S_0(u)$  has been measured under the same conditions and the results are shown in Figure 8. The values are nearly constant near the wake centre with a mean value of  $-0.36$ , but become rather indeterminate near the edge of the wake, where the experimental difficulty of measuring skewness in an intermittently turbulent flow is considerable. It was also confirmed that, within the limits of experimental error,  $S_0(v) = S_0(w) = 0$ , a result predicted by local isotropy.



The predictions of the original theory of local isotropy are therefore only verified for the central part of the wake.

### V. DISCUSSION OF RESULTS

The results of these experiments may be summarized by saying that those predictions of the theory of local isotropy which concern all the quantities measured in these experiments are verified for the central part of the wake only. On the other hand, predictions concerning the relative magnitudes of mean squares of velocity derivatives are confirmed over the entire observable wake,

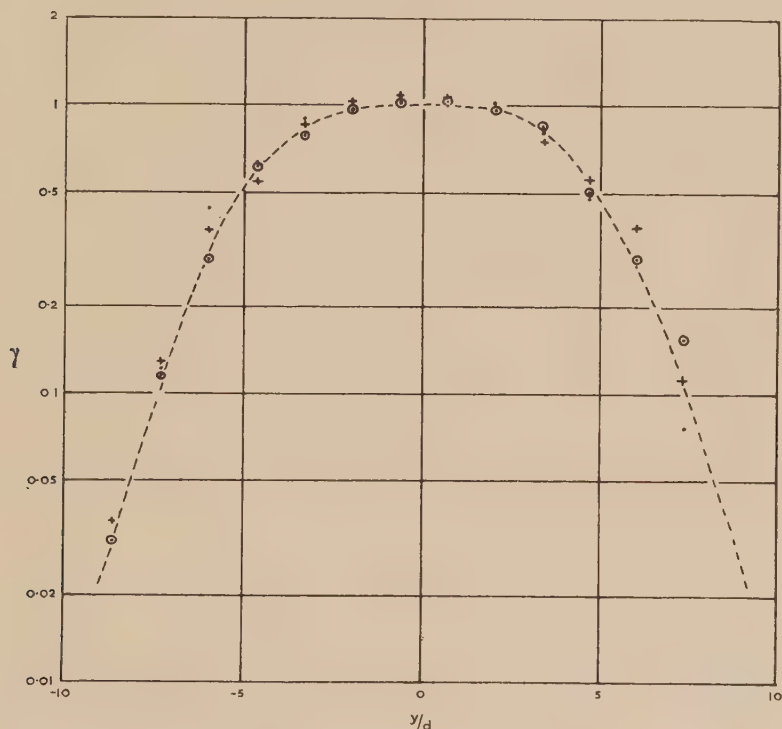


Fig. 10.—Variation of intermittency factor for  $x/d=120$

•	$T_0(u)$
+	$T_0(v)$
⊙	$T_0(w)$

although the predicted constancy of shape factors such as  $T_0(u)$  fails to appear. It is natural to assume that the observed intermittent nature of the turbulent flow is responsible for the failure of local isotropy in the outer portions of the wake, and it seems to be possible to explain the observed behaviour in terms of a modified local isotropy. It is observed that there is a relatively sharp division between regions of fully developed turbulence and regions of nearly laminar flow, and it may be postulated that the regions of fully developed turbulence have existed for a sufficiently long time to allow the establishment of local isotropy over substantially the whole region. The flow picture is of an intermittent

distribution of turbulence with the boundaries between regions of laminar and turbulent motion sharp and distinct, and of each individual region of turbulent motion having established local isotropy within its boundaries. Then, using  $\gamma$  to denote the mean fractional duration of turbulent flow at any given point, it is readily seen that the assumption of local isotropy within the turbulent regions leads to the relations

$$\overline{\left(\frac{\partial u}{\partial x}\right)^2} = \frac{1}{2} \overline{\left(\frac{\partial v}{\partial x}\right)^2} = \frac{1}{2} \overline{\left(\frac{\partial w}{\partial x}\right)^2}$$

$$S_0(u) = \gamma^{-\frac{1}{2}} S'_0(u)$$

$$T_0(u) = \gamma^{-1} T'_0(u)$$

$$T_0(v) = T_0(w) = \gamma^{-1} T'_0(v)$$

where  $S'_0(u)$ ,  $T'_0(u)$ , and  $T'_0(v)$  are the values obtained in steady homogeneous turbulent flow, for example, isotropic turbulence behind a grid.

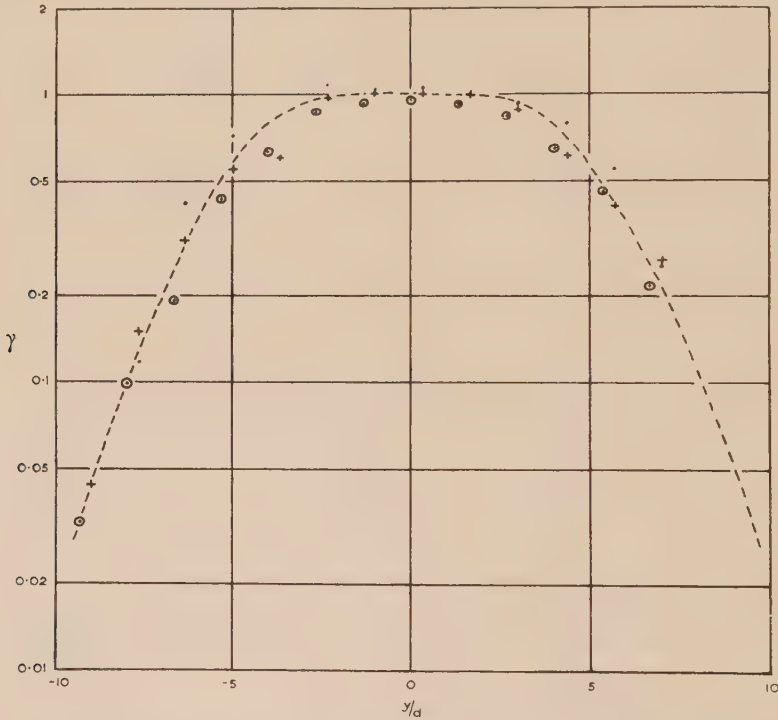


Fig. 11.—Variation of intermittency factor for  $x/d=160$

•	$T_0(u)$
+	$T_0(v)$
○	$T_0(w)$

Accepting this hypothesis, the flattening factor  $T_0(u)$  is a measure of the “gustiness” of the turbulence, and  $\gamma$  may be calculated from it. In Figures 9-11, values of  $\gamma$  calculated from observed values of  $T_0(u)$ ,  $T_0(v)$ , and  $T_0(w)$  are plotted, using values for  $T'_0(u)$  and  $T'_0(v)$  known from experiments on isotropic turbulence. The consistency of the values so obtained is good evidence of the truth of the hypothesis.

Towards the edge of the wake,  $\gamma$  becomes very small, as low as 0.04, which is roughly confirmed by the oscillograph observations. A similar structure of turbulence in a jet has been found by Corrsin(8), who explained his oscillographic observations of intermittency in the turbulent fluctuations as due to a phenomenon similar to transition in a laminar boundary layer, and classified the flow into a fully turbulent core, an intermittent "transition" region and a laminar collar. No evidence has been found here for the existence of any permanently laminar layer, and Corrsin's classification may well be a simplification of the observed continuous variation of  $\gamma$  from one to zero. The quantitative measurements given above show that this inhomogeneous structure is typical of the wake as a whole, and that except very close to the centre, the flow is essentially intermittent, the relative duration of the periods of laminar flow increasing rapidly with distance from the wake centre. The explanation of the intermittent nature of the turbulent motion must be sought in the processes whereby undisturbed fluid is entrained into the wake. The results show that this proceeds in short bursts of turbulence, which may be derived from initial disturbances that by virtue of their motion have been able to abstract energy from the mean flow, while lesser disturbances have grown much more slowly, if at all. Whatever the detailed explanation, the mechanism is likely to be common to all examples of free turbulence when there is mixing of undisturbed fluid with the turbulent flow.

The existence of local isotropy over the wake (in the sense described above) simplifies greatly the theoretical treatment of turbulent shear flow, and in particular makes possible simple and accurate measurements of the local energy dissipation due to viscosity. The energy dissipation equation becomes simply

$$\varepsilon = 15 \nu \overline{\left(\frac{\partial u}{\partial x}\right)^2}$$

and the distribution of dissipation across the wake is similar to that of  $\overline{\left(\frac{\partial u}{\partial x}\right)^2}$  (Figs. 2-4). It so becomes possible to investigate in detail the processes of transfer and diffusion of turbulent energy in the wake, and to form an estimate of the relative importance of pressure flow and diffusion in redistributing the turbulent energy. For this, measurements of the shearing stress, the energy diffusion components  $\overline{u^2v}$ ,  $\overline{v^3}$ , and  $\overline{w^2v}$ , and of the viscous dissipation are necessary. By difference, the pressure-velocity correlations may be evaluated and a complete account of the energy balance obtained. Information of this sort would be of considerable value in an attempt to apply the statistical theory of turbulence to shear flow.

In conclusion, the existence of local isotropy in the turbulent wake of a cylinder has been verified, but allowance must be made for the intermittent character of the flow. The flow alternates irregularly between laminar and turbulent flow but within each patch of turbulent flow, local isotropy obtains.

## VI. REFERENCES

- (1) KOLMOGOROFF, A. N.—*C.R. Acad. Sci. U.R.S.S.* **30** : 301-5 (1941).
- (2) KOLMOGOROFF, A. N.—*Ibid.* **32** : 16-18 (1941).
- (3) TOWNSEND, A. A.—*Proc. Camb. Phil. Soc.* in press.
- (4) TOWNSEND, A. A.—*Proc. Roy. Soc. A* **190** : 551-61 (1947).
- (5) BATCHELOR, G. K.—*Proc. Camb. Phil. Soc.* **43** : 533-59 (1947).
- (6) BATCHELOR, G. K., and TOWNSEND, A. A.—*Proc. Roy. Soc. A* **190** : 534-50 (1947).
- (7) TOWNSEND, A. A.—*Proc. Camb. Phil. Soc.* **43** : 560-70 (1947).
- (8) CORRISIN, S.—NACA Adv. Conf. Rep. 3L23 (1943).

# THE VELOCITY OF SOUND IN GASES

By R. L. ABBEY\* and G. E. BARLOW\*

[*Manuscript received December 2, 1947*]

## *Summary*

The velocity of sound in gases confined in a tube has been measured at a frequency of 1000 cycles per second using a method based upon acoustical feedback. Measurements of the velocity were made at pressures ranging from atmospheric down to several millimetres of mercury, for the following gases: air, nitrogen, oxygen, carbon dioxide, and methane. The values obtained for the velocity at atmospheric pressure agree very well with established figures, but the corrected velocity was found to vary with the pressure. The general validity of this tube correction is discussed.

## I. INTRODUCTION

This paper presents a determination of the velocity of sound in gases at low frequency and low pressures. The velocity was measured in a closed tube and the observed value was then corrected for the effect of the walls of the tube by means of the Helmholtz-Kirchhoff tube correction. According to this relation, the velocity of sound in a tube of diameter  $d$  at a frequency  $n$  c/s is less than the velocity  $V$  in the free gas by the amount  $\frac{Vc}{d\sqrt{\pi n}}$  where  $c$  is a function of the viscosity, thermal conductivity, specific heats, and density of the gas. This expression has been verified for a number of gases at atmospheric pressure over a range of values of  $n$  and  $d$ . Some workers have found the experimental value of  $c$  to be less than the theoretical, but for smooth-walled tubes this variation is small. It is negligible for the tube used in our work. Since  $c$  is inversely proportional to the square root of the density of the gas, it becomes large at low pressures and a more exact knowledge of this constant would be desirable. However, as previous workers found the errors in  $c$  are small, the general validity of the Helmholtz-Kirchhoff correction has been assumed in correcting our results. It is intended to study this correction more thoroughly later.

## II. METHOD

The method used is based upon the acoustical feedback between a microphone and loudspeaker which are connected to the input and output respectively of a tuned amplifier. Under certain conditions oscillations are produced (1, 2, 3).

\* Physics Department, University of Melbourne.



Referring to Figure 1, let the total amplification of the circuit for current of frequency  $\omega$  be  $A(\omega)$ , the phase change produced by the microphone, amplifier, and loudspeaker be  $\eta(\omega)$ , and the distance between the microphone and loudspeaker be  $d$ .

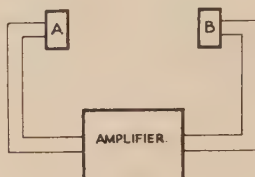


Fig. 1

In an equilibrium state, the change of phase produced in the whole circuit must be an integral multiple of  $2\pi$ . The change of phase in sound path  $AB$

$$= \frac{2\pi d}{\lambda}$$

$$= \frac{\omega d}{v}$$

since  $\lambda = \frac{2\pi v}{\omega}$  where  $\lambda$  is the wavelength and  $v$  the velocity of sound of frequency  $\omega$ ,

$$\therefore \eta(\omega) + \frac{\omega d}{v} = 2n\pi$$

$$d = \frac{v}{\omega} [2n\pi - \eta(\omega)] \quad \dots\dots\dots (1)$$

For a fixed value of  $d$ , equation (1) will in general have one solution  $\omega_n$  for  $\omega$  corresponding to each value of  $n$ .

Consider an initial time at which the currents corresponding to the frequencies  $\omega_n$  are  $i(\omega_n)$ . The times taken by each of these currents to traverse the circuit once are the same, say  $T$ . After time  $rT$  the corresponding currents will be

$$[A(\omega_n)]^r i(\omega_n)$$

and the total current

$$\sum_n [A(\omega_n)]^r i(\omega_n) = I_r.$$

Also as  $r \rightarrow \infty$ ,  $I_r$  tends to a finite value  $I$ .

$$I = \lim_{r \rightarrow \infty} \sum_n [A(\omega_n)]^r i(\omega_n).$$

If  $n_0$  is the value of  $n$  for which  $A(\omega_n)$  is greatest then,

$$I = \lim_{r \rightarrow \infty} [A(\omega_{n_0})]^r i(\omega_{n_0}) \left\{ 1 + \sum_{n \neq n_0} \left[ \frac{A(\omega_n)}{A(\omega_{n_0})} \right]^r \frac{i(\omega_n)}{i(\omega_{n_0})} \right\}$$

and since  $A(\omega_n) < A(\omega_{n_0})$ ,  $\sum_{n \neq n_0} \rightarrow 0$  as  $r \rightarrow \infty$  for any arbitrary set of values  $i(\omega_n)$ , provided  $i(\omega_{n_0}) \neq 0$ .

$$\therefore I = \lim_{r \rightarrow \infty} [A(\omega_{n_0})]^r i(\omega_{n_0})$$

i.e. the total current corresponds to only one allowed frequency—that for which the amplification of the circuit is greatest.

Consider now the effect of changing  $d$ . Suppose we start with  $\omega_n = \Omega$ , the resonant frequency, for  $n = n_1$ , by choosing a suitable distance  $d$ . Let  $d$  be increased. The frequency of the current will be equal to  $\omega_{n_1}$  which decreases with increasing  $d$ , until  $\omega_{n_1}$  is so low that  $A(\omega_{n_1+1}) > A(\omega_{n_1})$ . Then the current will change over to frequency  $\omega_{n_1+1}$ . This process may be repeated as  $d$  is still further increased, and the frequency will vary about the resonant frequency as in Figure 2. However,  $A(\omega)$  may be too small for any oscillation to occur in the circuit, and only the portion of the curve represented by the full line can be obtained. This frequency range is determined by the gain and selectivity of the amplifier, and amounted to about 10 c/s for our measurements in air at atmospheric pressure.

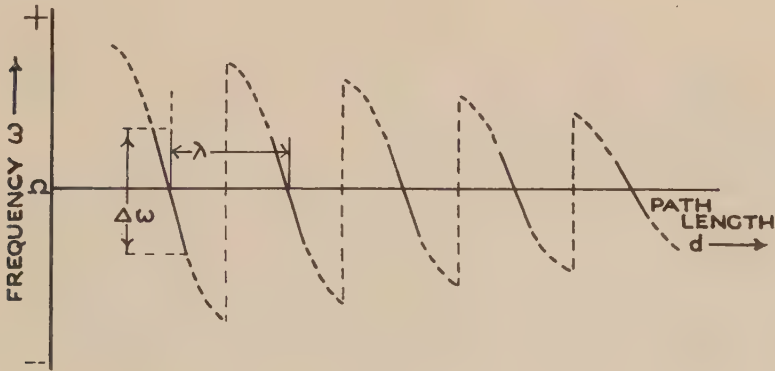


Fig. 2

The mean frequency recurs when the microphone is moved through one wavelength. Thus the effect provides a means of measuring wave speeds(3). The technique involves the measurement of the displacement of the microphone between two or more positions at which the frequency of oscillation is the same. It is to be noted that the method does not require the formation of stationary waves.

In the first experiments the speaker diaphragm was tuned to a frequency which could be conveniently compared with a standard. Since the measurement depends on the phase of feedback in the path AB as determined by the position of the microphone, any variation in phase of the signal due to a reactive change in the circuit is undesirable. A carefully stabilized amplifier containing a tuned circuit was used with a loudspeaker whose response peaked well outside the frequency range. The resonant frequency could thus be controlled by using a variable tuned circuit.

In order to control the conditions of temperature, pressure, water vapour, and carbon dioxide content of the gas, these measurements were made in a closed tube.

### III. APPARATUS

#### (a) General

The general layout of the apparatus is shown in Figure 3. The loudspeaker A is mounted at one end of a brass tube C, and the microphone B can slide along

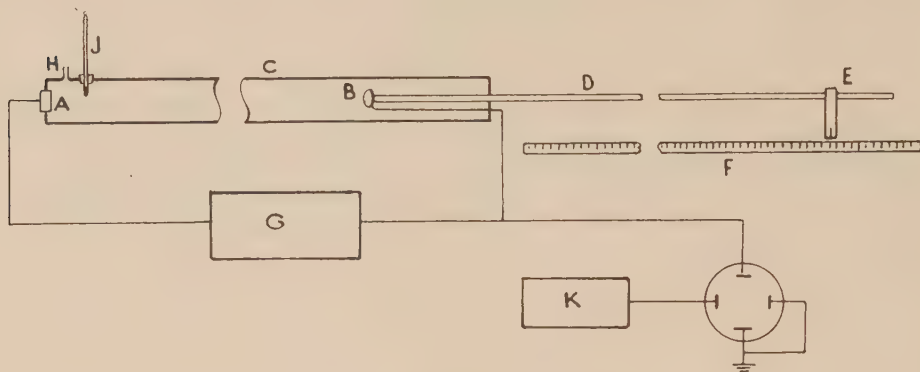


Fig. 3

the tube at the end of a rod D. The displacement of the microphone is measured by sighting a reading microscope E attached to the end of the rod, on to a graduated scale F. The microphone is connected to a speaker unit, through a tuned amplifier G. The amplifier output is also connected to the X deflecting plates of a cathode ray oscillograph. A standard frequency from the source K is applied to the Y plates and a frequency match can be obtained by means of the resulting Lissajous figure. The tube can be exhausted and filled with gas to the required pressure through the inlet H. A thermometer J measures the temperature of the gas.

#### (b) Tube

Details of the tube are shown in Figure 4. The tube is made of brass, 3 in. diameter, 6 ft. long, and having a wall thickness of  $\frac{1}{8}$  in. It is rigidly clamped to the bench so that the rod carrying the microphone moves parallel to the scale. This rod slides through a castor oil lubricated rubber tube L, and this provides a satisfactory vacuum seal, even at the lowest pressures used. The end of the rod bearing the reading microscope is attached to a solid runner on a guide rail and can be clamped and finally set by means of an adjusting screw. The displacement is measured on a 100-cm. brass scale which was calibrated against the standard metre in this laboratory. The scale is graduated in millimetres and the reading can be estimated to better than 0.1 mm.

The gas temperature is measured with a calibrated thermometer mounted about 50 cm. from the loudspeaker. A  $\frac{3}{4}$ -in. layer of hair-felt lagging around

the tube was sufficient to maintain a constant temperature, since the experiment was performed in a room where the change of temperature was less than  $0.5^{\circ}\text{C}$ . per hour, at certain times of the day.

(c) *Electrical System*

The speaker unit is a dynamic headphone with the response peaked high in the audio region. A carbon granule microphone, which was first tried, proved unsatisfactory owing to mechanical instability, and it was replaced by the crystal cartridge from a J.T. 30 model. The diaphragms of the microphone and loud-speaker were pierced with several holes to prevent damage as the pressure was changed. The circuit of the amplifier is shown in Figure 5. A high  $Q$  parallel resonant circuit is included in the anode circuit of the first tube, and can be tuned over a small frequency range about 1000 c/s by means of a midget variable capacitor. This is provided to correct changes in the peaked frequency due to

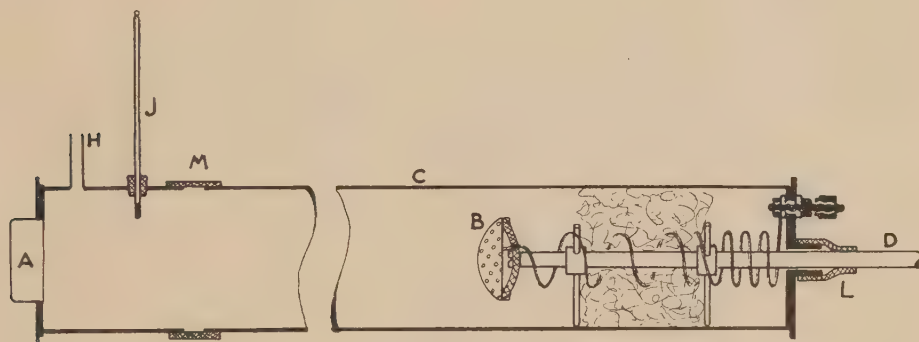


Fig. 4

temperature, humidity, etc. and to set this frequency at maximum intensity in the middle of the range of oscillation. The tuning can be sharpened to any desired degree by the positive feedback from  $V_3$  to  $V_1$ . Negative feedback provides stability of operation to line voltage fluctuations.

A 1000 c/s valve-maintained tuning fork oscillator is used as a frequency standard. This has a National Physical Laboratory certificate, and the calibration has been checked against the standard 1000 c/s signal provided by the P.M.G. Laboratory, the imprecision of the standard being less than 1 in  $10^6$ . All leads carrying audio frequency signals were shielded, and the apparatus was well earthed.

(d) *Vacuum System*

The gas outlet tube H is connected through chemical towers to either the pump or the gas supply as required, and also to a mercury manometer. All the joins of the tube are sealed with vacuum wax and the sliding rubber seal is lubricated with castor oil. This was sufficient to hold the pressure at less than 0.2 mm. Hg for over ten hours. Any leak was detected immediately by a change in frequency of the oscillations.

*(e) Special Precautions*

It was found in practice that a number of special precautions were necessary, principally to ensure that the only acoustical path from the speaker to the microphone was by way of the gas in the tube between them. The tube was cut near the speaker end and a reinforced rubber section M was cemented on, to leave a gap of about 1 cm. between the two sections of brass tube. This is efficient in reducing the conduction of sound along the metal walls, an effect which becomes serious at low pressures when the intensity of the gas-transmitted wave becomes small and the gain and sensitivity of the amplifier are used at the maximum. Under such conditions the apparatus was liable to break suddenly into oscillation of large amplitude, and the displacement of the microphone between two positions at which this effect occurred corresponded to the wavelength of sound in brass at a frequency of 1000 c/s.

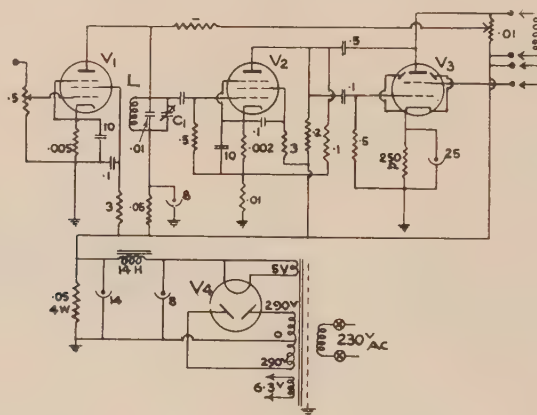


Fig. 5.—1 kc. tuned amplifier.  $L$ , 2.5 H air-cored inductor;  $C_1$ , 30  $\mu\text{F}$  midget variable capacitor;  $V_1$ ,  $V_2$ , 6J7G;  $V_3$ , 6Y6GT;  $V_4$ , 80; resistances are in megohms, capacities in microfarads.

It has been pointed out by Kaye and Sherratt(4) that the longitudinal vibration of the rod, caused by the intense sound waves impinging on the microphone and its holder, constitutes a possible source of error. We found that the vibration of the rod reacted on the microphone to a degree determined by the tightness of clamping the end of the former, and also by the length of the rod extending outside the tube. Kaye and Sherratt's method of eliminating this effect, in which they used a tightly fitting piston of heavy construction, was not desirable in our work because of the reflected wave forming a strong standing wave system. We tried several methods of mounting the microphone, such as on a plug of porous wood, but they were not satisfactory. Finally, the method shown in Figure 4 was adopted. The rod is centred in the tube by a pair of three-legged supports, which present a small area to the pressure wave. The microphone itself is mounted on a strip of thick rubber, screwed to the rod at its centre,



and cemented to the microphone at the ends. This effectively isolates the microphone from the rod. Its leads are then coiled in a spiral to connect with terminals on the end-plate of the tube.

As discussed by Henry(5), the vibration of the gas column behind the microphone is another source of error, and such vibrations are difficult to eliminate in a completely enclosed system. To reduce this effect, cotton wool was loosely packed between the two rod supports as an absorbing medium. It is essential that the absorbing material should not move with respect to the rod as the latter is moved backwards and forwards, or transmit vibrations to the rod. Solid porous material, which was first used, was rejected for this reason, and about 6 in. of loosely packed cotton wool was found to be quite satisfactory.

#### IV. MEASUREMENTS

Before an experiment was commenced the amplifier and associated electrical equipment were switched on and allowed to "warm up" for at least half an hour to reach stable operation. The amplifier was tuned by adjusting the trimming condenser  $C_2$  so that the maximum response at a frequency of exactly 1000 c/s was obtained at the centre of the range of oscillation. Operation off the centre frequency is indicated when successive wavelengths differ by more than the possible experimental error of the setting. The required sharpness of tuning was obtained by increasing the positive feedback until self-oscillation just occurred, and then reducing this feedback slightly. The gain control was then adjusted for a convenient size of image on the screen of the oscilloscope.

The tube was exhausted and filled with gas through a system maintained at a pressure greater than atmospheric. Evacuating and refilling were carried out slowly to avoid damage to the microphone and speaker by sudden changes in pressure, and also to reduce variations in the temperature of the gas caused by expansion. Sufficient time for steady conditions to be reached was allowed before readings were commenced.

Readings at pressures less than 20 mm. Hg were found to be very sensitive to small traces of gas impurities, and extremely high values of velocity were obtained at one stage of the measurements. These results were found to be caused by organic vapours from the waxes, cements, etc. used in the assembly of the tube components. Such values could not be reproduced when the tube was refilled with gas. For low pressure work the tube was allowed to remain under vacuum overnight, and it was evacuated and refilled several times before any measurements were made.

The microphone was then moved along the tube to a position where oscillation took place and adjusted until a stationary one-to-one Lissajous figure was formed on the screen of the oscilloscope. The scale was read, and the microphone then moved in the same direction to the next point in a similar manner. The readings were checked when moving the microphone in the opposite direction. Temperature and pressure were recorded before and after each set of readings. A further check was obtained by reversing the connections to the loudspeaker or microphone. This changes the electrical phase by  $\pi$  radians and displaces all the readings by half a wavelength.

*Tube Correction*

The velocity of sound in tubes differs from that in free air owing to the effect of the walls of the tube. The viscosity and thermal conductivity of the gas both lead to a reduction of the velocity in the tube below its value in free air. Helmholtz investigated this effect, and his work was later modified by Kirchhoff, resulting in the well-known Helmholtz-Kirchhoff tube correction formula

$$V_r = V \left( 1 - \frac{c}{2r(\pi n)^{\frac{1}{2}}} \right) \dots\dots\dots (2)$$

where  $V$  is the velocity of sound in free gas, at a frequency  $n$  c/s, and  $V_r$  is the corresponding velocity of sound in a tube of radius  $r$  cm.;  $c$  is the Kirchhoff constant, given by

$$c = \left( \frac{\eta}{\rho} \right)^{\frac{1}{2}} + \left( \frac{k}{\rho C_v \gamma} \right)^{\frac{1}{2}} (\gamma - 1) \dots\dots\dots (3)$$

where  $\eta$  is the viscosity of the gas

$\rho$  is the density of the gas

$k$  is the thermal conductivity of the gas

$\gamma$  is the ratio of specific heats of the gas

$C_v$  is the specific heat at constant volume of the gas.

Although some workers have criticized this tube correction formula, the weight of recent experimental and theoretical evidence favours its validity. In particular, Kaye and Sherratt(4) carried out an extensive experimental investigation on the subject. They concluded that the Helmholtz-Kirchhoff formula was correct in its statement of the influence of the tube diameter and the frequency on the velocity, but that the roughness of the surface of the wall had an appreciable effect. Later work, by Norton(6), confirms the validity of equation (2), but indicates that the value of  $c$  for smooth walled tubes is 10 per cent. less than the value given by equation (3). This agrees well with the results of Kaye and Sherratt.

The theoretical analysis of the problem was studied by Henry(5), who revised the derivation of the Helmholtz-Kirchhoff equation and also investigated the effect of factors which Kirchhoff had not taken into account. His conclusion was that the relation was theoretically sound and that the factors which Kirchhoff had neglected in his assumptions had a negligible effect on the velocity. As his paper was written before the work of Kaye and Sherratt, the discordant results of earlier workers remained unexplained.

From equation (3) it is seen that the tube correction will vary inversely as the square root of the density. Thus for low pressures this correction becomes very large. It is therefore to be expected that the tube velocity will correspondingly decrease if the corrected free gas velocity is independent of pressure. This effect does not appear to have been studied previously. Keesom(8) and others, who measured the velocity of sound at various pressures and at low temperatures, presented tube corrections which in several cases were independent of the pressure, and at other times varied in the correct manner. However, their measurements were not made at pressures low enough for the tube correction to become very large.

## V. RESULTS

*(a) Velocity of Sound in Air*

The results of our measurements are given in Table 1. We have given the tube velocity and the theoretical tube correction, together with the resulting "free gas" velocity. The velocity at the temperature of measurement was corrected to that at 20° C. by means of the relation  $V_t = V_0(1 + \alpha t)^{\frac{1}{2}}$  where  $\alpha$  has the ideal gas value of 0.00366. As all measurements were made within a range of 4°-5° C. about 20° C., no serious error is introduced by this method. The velocity at atmospheric pressure was reduced to 0° C. by this means and a small gas imperfection correction then applied, following the method of Hardy, Telfair, and Pielemeier(9).

The air was slowly passed through a tower containing potassium hydroxide pellets to remove carbon dioxide, and then through a phosphorus pentoxide tower to remove water vapour.

TABLE 1  
VELOCITY OF SOUND IN AIR AT 20° C.

Pressure (cm. Hg)	Tube Velocity (metres/sec.)	Tube Correction (metres/sec.)	Free Gas Velocity (metres/sec.)
76	342.9 <sub>3</sub>	0.50	343.4 <sub>3</sub>
40	343.0	0.69	343.7
15	342.7	1.13	343.8
3	345.3	2.52	347.8
2	345.0	3.08	348.1
0.5	346.2	6.17	352.4

$$V_0 = 331.4_0 \text{ m. sec.}^{-1}$$

Previous determinations of the velocity of sound in air show a wide variation in the values obtained by different workers. Kneser(10) considered that this was due to the effect of the vessel containing the gas, while Pielemeier(11) considered that the presence of carbon dioxide and water vapour in small proportions could cause the existing discrepancy.

Hardy, Telfair, and Pielemeier in 1941(9) made a close examination of the corrections necessary to be applied to measurements made under experimental conditions in order to reduce the results to standard conditions. The latter include temperature, pressure, frequency, carbon dioxide and water vapour content, and N<sub>2</sub>-O<sub>2</sub> composition. The day-to-day variation of the latter alone may cause an error of 0.017 per cent. They conclude that many of these corrections have either been ignored or incorrectly applied, and have corrected "all the important measurements" made since 1900 to standard conditions.

The weighted mean of these results is then  $V_0 = 331.45 \pm 0.05$  metres/sec. This value is in excellent agreement with their own experimental value of  $331.44 \pm 0.05$  metres/sec. and a recalculated value of  $331.46 \pm 0.05$  metres/sec. However, some measurements included in their table are still inconsistent, and the results of several reliable workers have not been considered, viz. Dixon, Campbell, and Parker(12) (331.8 metres/sec.); Cornish and Eastman(13) (331.5 metres/sec.); Kaye and Sherratt(4) (331.6 metres/sec.). Other determinations of the velocity of sound in air made since 1941 include the following: van Itterbeek and Vandoninck(14), 1944 (332.0 metres/sec.); Stewart(15), 1946 (331.7 metres/sec.); Bömmel(16), 1945 (331.5 metres/sec.).

The above value of  $V_0$  is in good agreement with that obtained by Hardy, Telfair, and Pielemeier. However, the increase in velocity below a pressure of 10 cm. Hg is much greater than any experimental error, although the value at 5 mm. Hg is likely to be in error because the sensitivity of the apparatus is not as great as at atmospheric pressure.

#### (b) *Velocity of Sound in Nitrogen*

Values of the velocity of sound in nitrogen, at various pressures, are given in Table 2. The gas used was commercial nitrogen, specified by the manufacturers as being 99.5 per cent. pure. It was dried by passing it through the KOH and  $P_2O_5$  towers as used for air.

TABLE 2  
VELOCITY OF SOUND IN NITROGEN AT 20° C.

Pressure (cm. Hg)	Tube Velocity (metres/sec.)	Tube Correction (metres/sec.)	Free Gas Velocity (metres/sec.)
76	348.3	0.50	348.8
50	347.7	0.62	348.4
25	348.0	0.87	348.9
15	347.8	1.12	348.9
10	347.6	1.38	349.0
5	347.6	1.94	349.5
2.5	347.2	2.76	350.0

$$V_0 = 336.64 \pm 0.16 \text{ m. sec.}^{-1}$$

The tube velocity decreases slightly, but as the tube correction is increasing, the resultant free gas velocity increases as the pressure is reduced, but to a much smaller degree than for air.



Other results for the velocity of sound in nitrogen are as follows :

	$V_0$
Keesom and van Lammeren(17) .. ..	337.1 metres/sec.
Dixon, Campbell and Parker(12) .. ..	337.5
Bömmel(16) .. ..	336.6
van Itterbeek and Mariëns(18) .. ..	337.5
Colwell and Gibson(19) .. ..	337.2

(c) *Velocity of Sound in Oxygen*

Commercial oxygen specified as being 99.0 per cent. pure was used, and passed through the KOH and  $P_2O_5$  towers as previously. Results are shown in Table 3.

TABLE 3  
VELOCITY OF SOUND IN OXYGEN AT 20° C.

Pressure (cm. Hg)	Tube Velocity (metres/sec.)	Tube Correction (metres/sec.)	Free Gas Velocity (metres/sec.)
76	325.7	0.47	326.2
50	325.7	0.58	326.3
25	325.9	0.81	326.7
15	325.6	1.06	326.6
10	325.8	1.29	327.1
5	325.4	1.82	327.2

$$V_0 = 314.84 \pm 0.12 \text{ m. sec.}^{-1}$$

Again the tube velocity is constant, so that there is an increase in the free gas velocity when the tube correction is applied.

Some of the previously reported results for the velocity of sound in oxygen are :

	$V_0$
van Itterbeek and Mariëns(18) .. ..	315.5 metres/sec.
Keesom, van Itterbeek, and van Lammeren(20) .. ..	315.4

(d) *Velocity of Sound in Carbon Dioxide*

Commercial carbon dioxide of a specified purity of 99.8 per cent. was used. Results are given in Table 4.

The results have not been reduced to 0° C. because of uncertainty in the temperature relation. However, our value for the velocity at 20° C. agrees with that of Kaye and Sherratt(4) which is 266.7 metres/sec. (from the reported value at 18° C.). Their value for the velocity reduced to 0° C. was 258.3 metres/sec., which compares well with other results at audible frequencies, e.g. Colwell and Gibson(19) 258.15 metres/sec.; Dixon, Campbell, and Parker(12) 258.3 metres/sec.



Again we find that the free gas velocity increases as the pressure is reduced, although the tube velocity increases slightly to a maximum at a pressure of about 20 mm. Hg and then decreases to a lower value at 3 mm. Hg.

TABLE 4  
VELOCITY OF SOUND IN CARBON DIOXIDE AT 20° C.

Pressure (cm. Hg)	Tube Velocity (metres/sec.)	Tube Correction (metres/sec.)	Free Gas Velocity (metres/sec.)
76	266.5	0.24	266.7
50	266.8	0.30	267.1
40	266.9	0.33	267.2
30	267.2	0.39	267.6
15	267.3	0.56	267.9
10	267.1	0.68	267.8
3	267.7	1.23	268.9

(e) *Velocity of Sound in Methane*

The gas used was commercial methane and the following analysis was supplied by the Chemistry Department, University of Melbourne: CH<sub>4</sub>, 95.5%; CO<sub>2</sub>, 1.6%; ethylene types, 0.6%; N<sub>2</sub>, 1.5%; O<sub>2</sub>, 0.5%. The CO<sub>2</sub> was removed by passing the gas through the KOH tower.

Results of the measurements are shown in Table 5.

TABLE 5  
VELOCITY OF SOUND IN METHANE AT 17° C.

Pressure (cm. Hg)	Tube Velocity (metres/sec.)	Tube Correction (metres/sec.)	Free Gas Velocity (metres/sec.)
76	439.9	0.60	440.5
50	440.2	0.74	440.9
25	440.1	1.06	441.2
15	439.3	1.36	440.7
10	439.0	1.66	440.7
5	438.8	2.34	441.1
3	436.9	3.02	439.9
1.5	434.3	4.27	438.5
1.0	433.6	5.23	438.8
0.6	435.2	6.75	442.0
0.5	433.2	7.40	440.6
0.3	433.8	9.56	443.4

As methane is of special interest because of the more complex structure of the molecule, and also because of the rather different variation of velocity with pressure which we observed, a larger number of readings were made over a period of several weeks. The results in Table 5 were quite reproducible within the limits of experimental error as indicated in the curves of Figure 7.

The velocity remains constant until a pressure of 5 cm. Hg is reached, when it decreases suddenly to a minimum at 1.5 cm. Hg, then increases rapidly at lower pressures.

## VI. ACCURACY OF RESULTS

As previously mentioned, two or three wavelengths in the tube could be measured for each gas used. A number of readings were obtained at each pressure, exhausting and refilling the tube before each determination. The average error is indicated on the curves. Although the sensitivity of the method is capable of giving an accuracy of less than 1 in 3000, the mean error is often greater than this, being about 1 in 1500. This is considered satisfactory, since the measurements were carried out over a period of six months, and repeated readings on any gas were made from time to time, using probably a fresh supply of the gas. Other factors, such as mechanical errors in alignment of the apparatus, would tend to reduce the ultimate accuracy of measurement.

## VII. CONCLUSION

Little information is available from other workers as to the variation of the velocity of sound with pressure, especially at low frequencies. Colwell and Gibson(19) determined the velocity of sound in air, carbon dioxide, and nitrogen, at pressures ranging from 176 cm. Hg to 26 cm. Hg, and reported no variation. However, the accuracy of their measurement was not high, and variations of more than 2 metres per second are shown in their results.

The work at the Leiden laboratory(8) was carried out at low temperatures obtained with liquid gases, ranging from less than 5° K. (liquid helium) in helium to about 160° K. (liquid ethylene) in oxygen. The gas is not perfect at these temperatures, and the velocity of sound depends on the pressure according to relations derived from the equation of state for the gas. The observations were carried out only down to 5 cm. Hg, and to obtain the velocity at zero pressure the pressure-velocity curve was extrapolated back to zero pressure, assuming that the curve continues in the same form below 5 cm. Hg. Our results indicate, however, that below 5 cm. Hg the velocity changes much more rapidly than is indicated by the simple extrapolation made by the Leiden workers. Furthermore, the Helmholtz-Kirchhoff tube correction was incorrectly used, because in Report No. 216 (c) of Keesom *et al.*(8) a constant tube correction of 0.2 metre per second was added to the tube velocity for each pressure.

Our results are shown in the curves of Figures 6 and 7, where the change in velocity from its value at atmospheric pressure is plotted against the pressure. In general, there is an increase of velocity as the pressure is reduced, this increase being more rapid at pressures below 7 cm. Hg. The velocity in air is 4 metres/sec.

greater at a pressure of 2.5 cm. Hg than at atmospheric pressure, although the velocity in nitrogen and oxygen has increased only slightly at this pressure. Carbon dioxide shows a continual increase from its value at atmospheric pressure

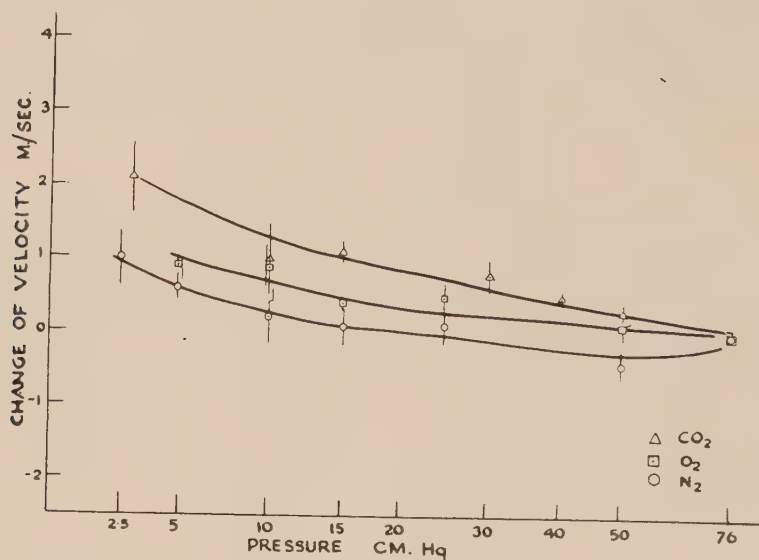


Fig. 6

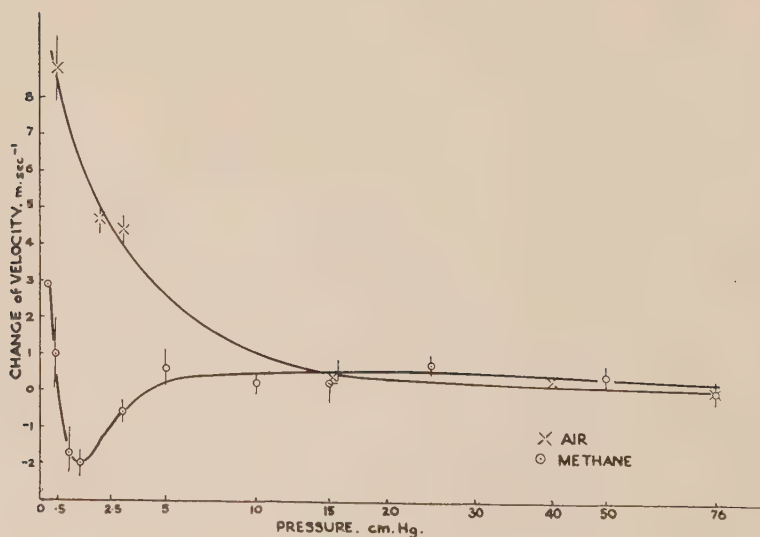


Fig. 7

to 2 metres/sec. greater than this at 3 cm. Hg. The behaviour of methane is anomalous, since the velocity decreases rapidly to a minimum at 1.5 cm. Hg, and then increases as the pressure is reduced still further.

The above results on the variation of the velocity with pressure were obtained from experimental measurements to which a theoretical tube correction has been added. As previously mentioned, the Helmholtz-Kirchhoff tube correction becomes very large at low pressures, so that the effect of any error in the formula becomes more serious as the pressure is reduced, and may account for the increase in velocity which was obtained. This may mean that a factor which varies with pressure was neglected in Kirchhoff's assumptions. The constants  $\eta$ ,  $k$ , and  $\gamma$  are theoretically independent of the pressure since the molecular mean free path is small compared with the dimensions of the apparatus, even at the lowest pressure used. A remaining possibility is that there is a real change in the free gas velocity of sound as the pressure is reduced. To study these effects it would be advantageous to use a frequency which is below the dispersion region, yet high enough for the tube correction to be negligible. It is proposed to extend the work along these lines.

### VIII. ACKNOWLEDGMENTS

The authors express their thanks to Professor L. H. Martin, University of Melbourne, for suggesting and supervising this research. One of us (G.E.B.) is indebted to the Dunlop Rubber Co. (Aust.) for a scholarship which enabled him to assist in this work.

### IX. REFERENCES

- (1) HUGHES, D. E.—*J. Soc. Electr. Engrs. Elec. (Lond.)* **12**: 245 (1883).
- (2) KENNELLY, A. E.—“Electrical Vibration Instruments”, p. 287 *et seq.* (Macmillan Co.: London, 1923.)
- (3) PLIMPTON, S. J.—*Amer. Phys. Teach.* **6**: 203 (1938).
- (4) KAYE, G. W. C., and SHERRATT, G. G.—*Proc. Roy. Soc. A* **141**: 123 (1933).
- (5) HENRY, P. S. H.—*Proc. Phys. Soc.* **43**: 340 (1931).
- (6) NORTON, G. A.—*J. Acoust. Soc. Amer.* **7**: 16 (1935).
- (7) FAY, R. D.—*Ibid.* **12**: 62 (1940).
- (8) KEESOM, W. H., *et al.*—*Leiden Comm.* **19** (1929-31), **20** (1931-33).
- (9) HARDY, H. C., TELFAIR, D., and PIELEMEIER, W. H.—*J. Acoust. Soc. Amer.* **13**: 226 (1942).
- (10) KNESER, H. O.—*Ann. Phys., Lpz.* **34**: 665 (1939).
- (11) PIELEMEIER, W. H.—*J. Acoust. Soc. Amer.* **10**: 313 (1939).
- (12) DIXON, H. B., CAMPBELL, C., and PARKER, A.—*Proc. Roy. Soc. A* **100**: 1 (1921).
- (13) CORNISH, R. E., and EASTMAN, E. D.—*Phys. Rev.* **33**: 258 (1929).
- (13) ITTERBEEK, A. VAN, and VANDONINCK, W.—*Ann. Phys., Paris* **19**: 88 (1944).
- (15) STEWART, J. L.—*Rev. Sci. Instrum.* **17**: 59 (1946).
- (16) BÖMMEL, H.—*Helv. Phys. Acta* **18**: 1 (1945).
- (17) KEESOM, W. H., and LAMMEREN, J. A. VAN—*Leiden Comm.* **10**: Rep. No. 221(c), p. 8 (1931-33).
- (18) ITTERBEEK, A. VAN, and MARIËNS, P.—*Physica Eindhoven* **4**: 207 (1937).
- (19) COLWELL, R. C., and GIBSON, L. H.—*J. Acoust. Soc. Amer.* **12**: 436 (1941).
- (20) KEESOM, W. H., ITTERBEEK, A. VAN, and LAMMEREN, J. A. VAN—*Leiden Comm.* **19**: Rep. No. 216(d) (1929-1931).

# THE DISORDERING OF $\beta$ BRASS BY COLD WORK

By R. W. K. HONEYCOMBE\* and W. BOAS\*

[*Manuscript received March 2, 1948*]

## Summary

The electrical resistivities of an  $\alpha$ - $\beta$  brass, some  $\alpha$  brasses of various zinc contents, and an aluminium bronze have been measured after various deformations by wire drawing. The resistivity of the duplex alloy increases steeply after about 80 per cent. reduction in area. This increase is shown to be due to the resistivity change of the  $\beta$  phase which probably becomes disordered on deformation. A recovery effect is observed in the duplex alloy on aging at room temperature, indicating that the disordered state is not stable. No recovery has been observed in single phase wires. Unusually large increases in resistivity of some  $\alpha$  phase alloys with deformation have been observed and may indicate that even these alloys are not completely disordered.

## I. INTRODUCTION

The  $\beta$ - $\beta'$  transformation in brass has been the subject of many investigations in recent years. In no case has the disordered state been preserved at room temperature by quenching. This is shown by the fact that the electrical resistivity of quenched  $\beta$  brass is very close to that of slowly cooled  $\beta$  brass. Smith(1) attempted to produce the disordered state by cold work, but reported that the resistivity changes less than 1 per cent. after wire drawing to 9.5 per cent. reduction in area. However, these experiments afford little possibility of producing the disordered structure because of the very small amount of deformation which could be imposed on the  $\beta$  brass before rupture.

The authors have shown that an  $\alpha$ - $\beta$  brass containing up to 80 per cent.  $\beta$  phase can be deformed to 90 per cent. reduction in area by wire drawing, during which process severe deformation occurs in both phases(2). The use of a duplex alloy in which crystals of a hard phase are embedded in crystals of a ductile phase may thus afford a general means of heavily deforming the hard phase without rupture. If in consequence of such a deformation the order in the alloy is destroyed, a marked increase in electrical resistivity occurs. For example, ordered single phase copper-gold and iron-nickel alloys have been cold worked and the establishment of the disordered state has been followed in this manner(3). In a brass where both kinds of atoms have nearly the same scattering factor for X-rays, the measurement of the electrical resistivity is the most convenient method for obtaining evidence of the ordered state.

This paper describes the changes in specific resistivity of  $\alpha$ - $\beta$  brass wires deformed in an attempt to produce the disordered state in the  $\beta$  phase. Some single phase copper alloys have also been examined.

\* Division of Tribophysics, C.S.I.R.



## II. EXPERIMENTAL

The materials used in the experiments were an  $\alpha$ - $\beta$  brass wire, 0.187 in. in diameter, and wires of single phase  $\alpha$  solid solutions 0.192 in. in diameter. The chemical analyses are contained in Table 1.

All the wires were annealed for 30 minutes at 800° C. and then slowly cooled in the furnace, taking approximately 12 hours to reach room temperature.

The wires were then drawn through a set of carbide dies using a mechanized wire drawing machine. The speed of drawing was about 4 ft./min. and stearic acid was used as a lubricant.

The electrical resistivity was measured within two hours of drawing on 20 cm. lengths of wire using a Kelvin double bridge. The temperature in each series was constant to within  $\pm 2.5^\circ$  C. and the accuracy of the results was about 0.5 per cent.

TABLE 1  
COMPOSITIONS OF COPPER ALLOYS USED IN THE INVESTIGATION

	Brass 1 ( $\alpha$ - $\beta$ )	Brass 2 ( $\alpha$ )	Brass 3 ( $\alpha$ )	Brass 4 ( $\alpha$ )	Aluminium Bronze ( $\alpha$ )
	%	%	%	%	%
Copper .. ..	58.5	64.8	79.75	91.36	93.95
Zinc .. ..	41.3*	35.1*	20.21*	8.59*	0.28
Aluminium .. ..					5.67*
Iron .. ..	0.13	0.02	0.02	Trace	0.1
Lead .. ..	0.07	0.03	0.01	Trace	
Tin .. ..	Nil	0.01	0.01	Trace	
Nickel .. ..	Nil	0.01			
Others .. ..			Trace	0.05	

\* By difference.

## III. RESULTS

## (a) Duplex Brass Wires

After the annealing treatment the  $\alpha$ - $\beta$  brass wires consisted of a very coarse distribution of  $\alpha$  crystals in the  $\beta$  matrix, the phases being present in approximately equal proportions. Their specific resistivities after various reductions in area by drawing are shown in Figure 1A. The specific resistivity increases almost linearly at first, but more steeply at higher deformations. After 95 per cent. reduction in area the specific resistivity had changed from  $6 \times 10^{-6}$  to  $8 \times 10^{-6}$  ohm-cm., an increase of 33 per cent. On further drawing to 98.1 per cent. reduction in area the resistivity decreased slightly.

Measurements made one day after drawing showed (Fig. 2) that the resistivity decreased with time. After six days, in the more heavily deformed wires, this recovery was between 12 and 25 per cent. of the total change in resistivity produced by cold drawing. Recovery had almost ceased after 27 days; the further decrease in resistivity after this period up to 79 days was just above the limit of accuracy of the measurements.

After aging for 42 days, one set of wires was annealed in hydrogen for 30 minutes at  $800^{\circ}\text{C}$ . and furnace cooled, i.e. the wires were subjected to the same treatment as they had received prior to drawing. Thus no difference in the

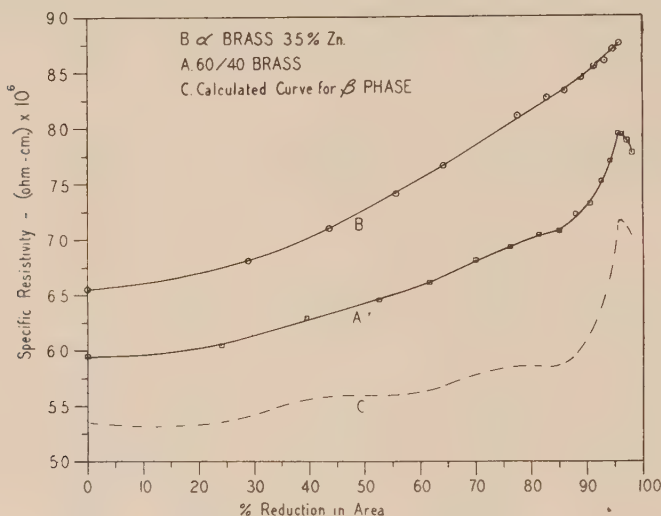


Fig. 1.—Specific resistivity of cold drawn  $\alpha$  and duplex brass wires.

distribution of the two phases should have occurred. The results of resistivity measurements on these wires are also plotted in Figure 2. This shows that all the annealed wires have approximately the same resistivity whatever the amount of previous cold work.

#### (b) Single Phase Wires of 35 Per Cent. Zinc

To separate the behaviour of the  $\alpha$  and  $\beta$  phases in the duplex wire, resistivity measurements were made on a single phase  $\alpha$  brass wire (35 per cent. zinc), the composition of which approximated to that of the  $\alpha$  phase in the duplex brass. The measurements showed that the specific resistivity of the  $\alpha$  brass wire also increased markedly with deformation, 90 per cent. reduction in area producing a change from  $6.5 \times 10^{-6}$  to  $8.5 \times 10^{-6}$  ohm-cm. (Fig. 1B). However, here the increase was uniform up to high deformations.

This is in contrast to the behaviour of the duplex brass wires, as is also the fact that practically no recovery occurs. This was established by measurements made on the single phase wires after periods of one and six days after

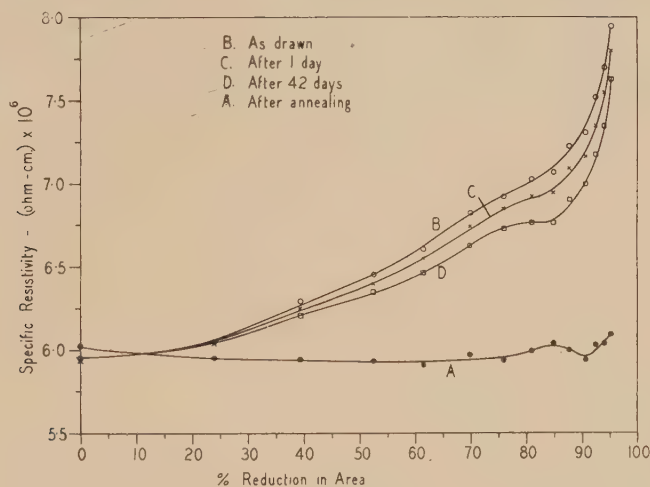


Fig. 2.—Effect of recovery on the specific resistivity of cold drawn duplex brass wires.

drawing to different degrees (Table 2). Although the changes are of the order of the experimental error it should be noted that in no case is there an increase in resistivity with time.

TABLE 2  
SPECIFIC RESISTIVITY OF 35 PER CENT. ZINC BRASS

Per Cent. Reduction in Area	Specific Resistivity (ohm-cm.)		
	Immediately after Drawing	One Day after Drawing	Six Days after Drawing
0	$6.55 \times 10^{-6}$	$6.55 \times 10^{-6}$	$6.54 \times 10^{-6}$
55.6	7.41	7.41	7.39
64.1	7.66	7.65	7.63
77.6	8.10	8.10	8.08
82.3	8.27	8.26	8.23
85.8	8.32	8.31	8.29
88.6	8.44	8.42	8.39
91.1	8.54	8.53	8.49
93.0	8.59	8.56	8.50
94.5	8.70	8.67	8.64
95.6	8.75	8.70	8.66

*(c) Single Phase Wires of Various Compositions*

The behaviour of the 35 per cent. zinc brass wire in the above experiments was contrary to that expected from the literature. It is generally thought that severe cold work changes the specific resistivity of most pure metals by less than 5 per cent. and that solid solutions behave in a similar manner(4). However, the above results indicate that the specific resistivity of brass containing 35 per cent. zinc has increased by about 30 per cent. after 90 per cent. reduction in area.

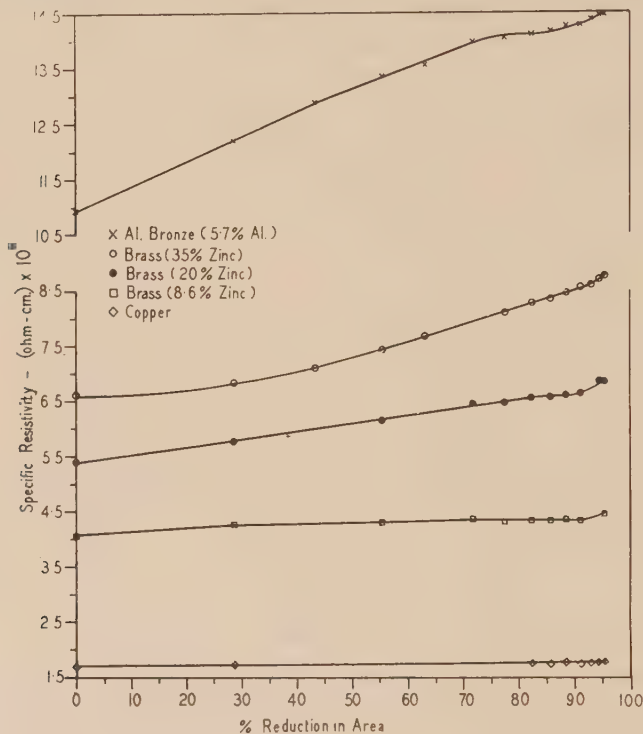


Fig. 3.—Specific resistivity of cold drawn copper base alloys.

To determine whether copper alloys of other compositions behave similarly, the brasses 3 and 4 and an aluminium bronze were investigated. As a check, pure copper wire was also examined.

The results obtained are plotted in Figure 3. It is clear that pure copper undergoes only a slight increase in electrical resistivity, the change amounting to only 3 per cent. after 90 per cent. reduction in area. With the 8.6 per cent. zinc brass, again only a small increase in resistivity occurred, the change after 90 per cent. reduction being approximately 6 per cent. However, an increase of 22 per cent. was observed in the 20 per cent. zinc wire after 90 per cent. reduction in area and of 30 per cent. in the aluminium bronze. Thus the effect is not specific to the addition of zinc to the copper.

The measurements were repeated on the 20 per cent. zinc alloy after 12 days and as in the 35 per cent. zinc alloy, the recovery observed was very small.

## IV. DISCUSSION

All the annealed duplex brass wires have nearly the same resistivity. This fact alone is not sufficient to indicate that the increase in resistivity on drawing is not due to the formation of cracks. However, a microscopic examination of sections of the drawn wires failed to reveal any cracks. Further, it is most unlikely that the wires could have been drawn to 98.1 per cent. reduction if cracks had been present previously, and, moreover, in this case the resistivity would have increased and not decreased during the last stages of drawing.

The behaviour of a wire of the pure  $\beta$  phase can be deduced from the curves (Figs. 1A and 1B) if it is assumed that:

- (1) The alloy consists of equal parts by volume of each phase.
- (2) The crystals in the duplex alloy change their resistivities with deformation in the same way as in the single phase alloys.
- (3) The deformation of the  $\alpha$  phase in the duplex alloy is the same as that of this alloy.

Curve C (Fig. 1) has been calculated using these assumptions.

Whereas the first assumption is borne out by microscopic observation, and the second by the success of the rule of mixtures applied to such alloys, indicating that the interaction between crystals is negligible as far as the electrical resistivity is concerned, the last assumption is certainly not correct. We have shown earlier(2) that the  $\alpha$  phase in the duplex brass is always deformed to a larger extent than the  $\beta$  phase, i.e. its deformation is larger than indicated by the total deformation of the alloy. Hence the curve B should be compressed horizontally at small and medium deformations in order to show the change in resistivity of the  $\alpha$  phase against the total deformation of the duplex alloy. On the other hand, the curve B is more nearly correct at high deformations since both phases then deform approximately to the same extent. Hence if the unequal deformation of the  $\alpha$  and  $\beta$  phases is taken into account, the curve C of the resistivity of the  $\beta$  phase will be lowered at small and medium deformations whereas at high deformations the sharp increase will be maintained. Thus this sharp rise will be still more pronounced than in the plotted curve.

As an explanation of this behaviour it is suggested that the  $\beta$  phase becomes disordered as a result of the cold working. This occurs only after an amount of deformation apparently larger than in other ordered alloys; however, the actual deformation suffered by the  $\beta$  phase is smaller than indicated by the total deformation of the duplex alloy. No definite explanation of the decrease in resistivity at very high reductions will be offered. Perhaps the heat developed during drawing produces a partial annealing of the wires. The fact that recovery occurs in the duplex but not in the single phase  $\alpha$  wires shows that the disordered state into which the  $\beta$  phase is brought by deformation cannot be stable at room temperature.

The change of resistivity with deformation of copper and an 8.6 per cent. zinc brass is as small as reported in the literature. However, the resistivities of 20 and 35 per cent. zinc brasses and of the aluminium bronze increase considerably. There is, of course, the possibility of the existence of a certain degree of



order also in those alloys. Guinier(5) has found that perfect disorder in a solid solution does not exist but that there is a certain extent of short range order. It is not clear whether such a short range order affects the electrical resistivity. According to Nix and Shockley(4) local short distance order does not affect such properties as electrical resistivity, and Muto(6) has calculated the resistivity as a function of the long range order parameter. On the other hand, similar claims had been made by Linde(7) and Muto regarding the insensitivity of X-ray interferences to order, if the coherent regions are small and irregular, before Guinier showed the effect experimentally. The possibility of the short range order affecting the electrical resistivity can therefore not be excluded.

#### V. ACKNOWLEDGMENTS

This work forms part of the research programme of the Division of Tribophysics, Council for Scientific and Industrial Research, Australia. We wish to thank Dr. S. H. Bastow, Chief of the Division, for his interest in the work and for helpful discussions. We are grateful to Professor E. J. Hartung for laboratory facilities in the Chemistry School, University of Melbourne; to Professor H. K. Worner for use of equipment in the Rosenhain Memorial Laboratory; and to the staff of the Munitions Supply Laboratories, Maribyrnong, for carrying out spectrographic analyses.

#### VI. REFERENCES

- (1) SMITH, C. S.—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **152**: 144 (1943).
- (2) HONEYCOMBE, R. W. K., and BOAS, W.—*Nature* **159**: 847 (1947); *Aust. J. Sci. Res. A* **1**: 70 (1948).
- (3) DAHL, O.—*Z. Metallk.* **28**: 133 (1936).
- (4) NIX, F. C., and SHOCKLEY, W.—*Rev. Mod. Phys.* **10**: 57 (1938).
- (5) GUINIER, A. J.—*Proc. Phys. Soc.* **57**: 310 (1945).
- (6) MUTO, T.—*Sci. Pap. Inst. Phys. Chem. Res. Tokyo* **30**: 99 (1936).
- (7) LINDE, J. O.—*Ann. Phys., Lpz.* **30**: 151 (1937).

# THE RATE OF EVAPORATION OF WATER THROUGH DUPLEX FILMS

By A. R. GILBY\* and E. HEYMANN\*

[*Manuscript received March 12, 1948*]

## *Summary*

The stability of duplex films of hydrocarbon oils containing spreaders is discussed in this paper. Duplex films containing spreaders of high molecular weight and of complex nature, e.g. stand oil, eucalyptus residue, polymerized oleic acid, may remain spread for long periods but are thermodynamically metastable.

The penetration of water molecules from a water surface through a duplex film has been discussed in terms of the evaporation resistances of the interfacial layer of spreader, the bulk layer of oil and the stagnant layer of air above the film.

Evaporation through duplex films of thickness 1–100  $\mu$  has been studied in conditions of "still air" at atmospheric pressure and in a vacuum, and in non-turbulent winds at several controlled wind velocities.

The efficiency of duplex films in reducing evaporation is greater in wind than in still air and increases with the wind velocity.

When duplex films thicker than 10  $\mu$  are placed on water surfaces, even a wind at 8 miles per hour does not increase the rate of evaporation. Thus with such films the resistance of the stagnant layer of air above the film is small compared with that of the oil layer and of the interfacial film of spreader.

With all duplex films thicker than 10  $\mu$  the total evaporation resistance is proportional to the film thickness and depends on the nature of the spreader. An analysis of the experimental results suggests that the interfacial layer of spreader makes an appreciable contribution to the evaporation resistance of the duplex films under investigation. Approximate relative figures characterizing the magnitude of this contribution are calculated for several spreaders.

## I. INTRODUCTION

Rideal(1) and Langmuir(2) have shown that the retarding effect of monolayers on the rate of evaporation from a water surface is noticeable only when evaporation into a vacuum is studied, the effect of monolayers being small compared with that of the stagnant layer of air above the surface under atmospheric conditions. Sebba and Briscoe(3), on the other hand, found that the resistance which monolayers offer to evaporation is appreciably increased on compression. Langmuir and Schaefer(4) have made a detailed study, both theoretical and experimental, of the influence of compressed monolayers and also of some duplex films on the rate of evaporation.

\* Chemistry Department, University of Melbourne.

A duplex film is a multimolecular film which is thick enough for the film-forming substance to have the physical properties it has in bulk and yet thin enough for the effect of gravity to be neglected. It may be obtained by spreading a hydrocarbon oil by the aid of suitable amphipathic compounds, "spreaders", and consists of a comparatively thick (several  $\mu$ ) layer of hydrocarbon oil containing some spreader, and an interfacial layer of spreader.

Heymann and Yoffe(5) found that duplex films of thickness about  $5\mu$  and consisting of paraffin oil containing spreaders of high molecular weight (*vide infra*) may reduce the evaporation of water to 15 per cent. of the original value. They made experimental and theoretical studies of these films with a view to conserving water reservoirs in arid areas, but reached the conclusion that such films are metastable and revert, though often after long periods of time, to lenses in equilibrium with a monolayer of spreader.

## II. CONDITIONS OF INITIAL SPREADING AND THE FINAL EQUILIBRIUM

We owe to Harkins(6) a general thermodynamic theory of spreading and can, therefore, limit ourselves to the discussion of duplex films of heavy hydrocarbon oils containing spreaders.

The decrease in free surface energy during the spreading of an oil, or more generally a non-miscible organic liquid, on water is

$$-\Delta G_S = \gamma_w - \gamma_o - \gamma_{wo} = F_S \quad \dots\dots\dots (1)$$

where  $\gamma_w$  is the surface tension of the water,  $\gamma_o$  the surface tension of the oil, and  $\gamma_{wo}$  the interfacial tension water-oil.  $F_S$  is called the initial spreading coefficient. Moreover,

$$F_S = W_A - W_C, \quad \dots\dots\dots (2)$$

where  $W_A = \gamma_w + \gamma_o - \gamma_{wo}$  is the work of adhesion between water and oil, and  $W_C = 2\gamma_o$  is the work of cohesion of the oil. If  $F_S < 0$  the oil will remain as a lens; if  $F_S > 0$  spreading will occur, and if  $F_S = 0$  an infinite lens of lens angle zero will result.  $F_S = 0$  is thus the limiting condition for spreading.

For a heavy hydrocarbon oil  $F_S < 0$ . Such oils therefore remain as lenses on a water surface.

$F_S$  can be made positive by reducing  $\gamma_{wo}$ . This may be achieved by dissolving in the oil an amphipathic compound, such as a fatty acid or a fatty alcohol, which is adsorbed in the interface between oil and water. In this way,  $\gamma_{wo}$  is reduced to  $\gamma'_{wo}$ , whilst  $\gamma_o$  is not strongly affected.

Thus  $\gamma'_{wo} < \gamma_{wo}$  and  $\gamma'_o \approx \gamma_o$ .

If the concentration of the amphipathic compound exceeds a certain, usually rather small, value

$$F_S = \gamma_w - \gamma'_o - \gamma'_{wo} > 0 \quad \dots\dots\dots (3)$$

As a consequence, spreading will occur. It may be regarded as being due to the expanding tendency of the film of the spreader in the interface, the lateral interfacial pressure of this film,  $\pi_i = \gamma_{wo} - \gamma'_{wo}$ , counteracting the contracting tendency of the lens.

On spreading, the whole or part of the available water surface may be covered and a duplex film may be formed. Such duplex films may persist for some time, but eventually they break up and a final state is reached in which lenses of oil containing some spreader are in equilibrium with a monolayer of spreader. The attainment of this equilibrium may be visualized in two ways:

(a) The hydrocarbon phase will collect to lenses because the work of cohesion of this phase (containing dissolved spreader),  $W'_C$ , is greater than the work of adhesion,  $W'_A$ , between the hydrocarbon phase and the upper part of the interfacial spreader film consisting essentially of hydrocarbon chains. Heymann and Yoffe(7) found that  $W'_A - W'_C$  is of the order of  $-1$  to  $-5$  erg cm.<sup>-2</sup> and attempted a theoretical explanation of this fact.

(b) Another mechanism for the attainment of the final equilibrium is suggested by an interesting observation due to Mercer(8) who showed that, during the spreading of lenses of paraffin oil containing stearic acid on alkaline solutions of calcium chloride, stearic acid molecules escape from the interface of the oil lens into the free water surface ahead of the spreading lens. A monolayer is thus built up in the water surface. The surface pressure  $\pi = \gamma_w - \gamma'_w$  of this monolayer counteracts the spreading pressure of the lens containing spreader. Eventually this surface pressure becomes sufficiently great to outweigh the spreading pressure of the lens, which has somewhat diminished by the loss of spreader molecules from the interface to the surface. As a result, the duplex film is pushed back into one or several lenses in equilibrium with a monolayer(5).

This equilibrium is characterized by the "final" or "equilibrium" spreading coefficient

$$F'_S = \gamma'_w - \gamma'_0 - \gamma'_{w0} = W'_A - W'_C < 0 \quad \dots\dots\dots (4)$$

where  $\gamma'_w$  is the surface tension of the water covered by a film of the spreader in equilibrium with lenses of the oil phase. All quantities in equation (4) refer to systems in which the phases are mutually saturated.

In the last stages of the process leading to equilibrium, the expanding monolayer of spreader pushes back the duplex film, or rather the extended lens, into one or several lenses with finite lens angles. This is a spontaneous process and therefore accompanied by a decrease of free surface energy. At the same time, however, gravitational potential energy is generated as more oil is piled up into the lenses. Equilibrium is reached when the rate of decrease of free surface energy becomes equal to the rate of increase of gravitational potential energy. The equilibrium corresponds to a definite lens angle from which  $F'_S$  may be calculated. This method of determining  $F'_S$ , due to Miller(9), is probably more accurate than that based on separate determinations of  $\gamma'_w$ ,  $\gamma'_0$  and  $\gamma'_{w0}$ , used by Heymann and Yoffe, because of cumulative errors of the latter method.

That the final equilibrium is one of lenses in equilibrium with a monolayer, has been noticed by several investigators. With none of the simple amphipathic compounds as spreaders, viz. palmitic acid, oleic acid, ricinoleic acid, linoleic



acid, are stable duplex films obtained\* and  $F'_s$  is always negative (Heymann and Yoffe).

The persistence of a duplex film over an appreciable period may be expected if the final equilibrium is reached slowly. Basing our discussion on mechanism (b) we may expect this to occur when the rate at which spreader molecules are shed from the interface into the surface is small. In this case a non-equilibrium state may persist for a considerable time during which  $\gamma'_{w_0}$  has reached a low value, which may be near its equilibrium value, whilst the monolayer in the water surface is still far from its equilibrium value, the surface tension of the water containing some spreader being somewhat below  $\gamma_w$ , but still appreciably above  $\gamma'_w$ . At this stage the spreading coefficient may remain positive over a period.

A slow attainment of equilibrium due to a small rate of shedding may be expected with spreaders of high molecular weight and complex structure. In fact, Heymann and Yoffe found that duplex films lasting for many months were obtained on using, as spreaders, substances such as polymerized oleic, ricinoleic and linoleic acids, stand oil, linseed oil, distillation residues of eucalyptus amygdalina oil, bitumen and stearin pitch. Similar results were obtained with paraffin oils irradiated by ultra-violet light in presence of air, with lubricating oils heated to 200° C. (Langmuir and Schaefer, 10), and with high boiling fractions (250-300° C.) and petroleum ether extracts of the distillation residue of the neutral oil of vertical retort tar (A. R. Docking,† K. N. Mortenson†).

The small rate of shedding from the oil-water interface into the water surface is due, in these cases, to the high viscosity and rigidity of the interfacial spreader film between oil and water as shown by ultra-microscopic observation of particles present in the interfacial layer (Heymann and Yoffe). The films develop high viscosity after hours and rigidity after days. Once a rigid skin is formed, the processes leading to the final equilibrium are very much retarded. This probably accounts for the long life of some metastable duplex films.

Dust particles, containing organic matter, and also mineral dust of a porous nature, destroy such films within several days, the dust particles penetrating the rigid interfacial layer and leaving innumerable tiny holes. Thus, whilst duplex films may well be a suitable means of killing mosquito larvæ because this action is comparatively fast, they are not practicable for the purpose of reducing evaporation from dams and reservoirs.

### III. THE RATE OF PENETRATION OF WATER THROUGH A DUPLEX FILM (THEORY)

It appears to be simplest to regard the rate of evaporation through a duplex film as governed by a diffusion process through a layer of oil of thickness  $d'$ , below which there is water of a fugacity corresponding to that of saturated water vapour ( $f_0$ ) and above which there is water vapour at a fugacity ( $f$ ) corresponding

\* The same applies to heavy metal soaps (e.g. copper oleate) and many commercial detergents, mostly of U.S.A. origin.

† Unpublished results from this laboratory.



to that of the humidity of the air (which is assumed to be constant). If there is no convection in the oil layer, the problem is very similar to that of diffusion of a gas through a porous plate, or a solute through a membrane.\* In this case a steady state of flow is set up, i.e. as much gas or solute leaves a given volume element as enters it per unit time(11). For this case, Fick's law gives

$$Q = -D' \frac{\partial f}{\partial d'} \dots \dots \dots (5)$$

where  $Q$  is the rate of penetration in the steady state of flow through unit area,  $\frac{\partial f}{\partial d'}$  is the fugacity gradient of water across the oil film, and  $D'$  is the diffusion coefficient. The solution of equation (5) is

$$Q = D' \frac{f_0 - f}{d'} \dots \dots \dots (6)$$

or

$$Q = \frac{f_0 - f}{d'/D'} = \frac{f_0 - f}{R'} \dots \dots \dots (7)$$

where  $R' = d'/D'$  is the diffusion resistance due to the oil layer.

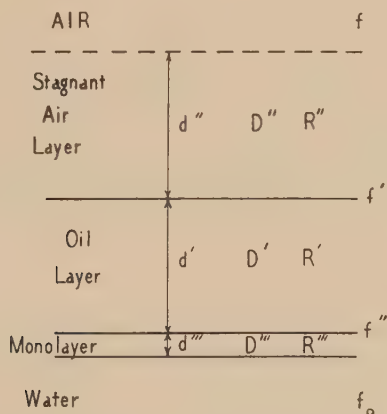


Fig. 1.—Model of evaporation of water through a duplex film.

Account must not only be taken of the resistance to evaporation of the oil layer, but of the interfacial film of spreader between oil and water ("monolayer") and of the stagnant layer of air above the film surface. If we consider evaporation as diffusion through these three layers in series, we obtain the model shown in Figure 1.

Let the fugacity of water in the bulk phase of water be  $f_0$ , that at the boundary between the interfacial film of spreader and the oil layer  $f''$ , at the surface of the oil layer  $f'$ , and in the air above the stagnant air layer  $f$ ;  $d'''$ ,  $d'$ ,

\* Mathematically this problem is identical with that of the flow of heat in a solid medium resulting from an unequal temperature distribution.

and  $d''$  are the thicknesses of the interfacial film (monolayer), oil layer, and stagnant air layer, and  $D'''$ ,  $D'$ , and  $D''$  are the corresponding diffusion constants of water in these layers. Now assuming the fugacity gradient in each layer to be constant and characteristic for that layer, we obtain for the steady state

$$Q = D''' \frac{f_0 - f''}{d'''} = D' \frac{f'' - f'}{d'} = D'' \frac{f' - f}{d''} \quad \dots\dots\dots (8)$$

Now since  $(f_0 - f'') + (f'' - f') + (f' - f) = f_0 - f$ ,

$$f_0 - f = Q \left( \frac{d'''}{D'''} + \frac{d'}{D'} + \frac{d''}{D''} \right)$$

or

$$Q = \frac{f_0 - f}{\frac{d'''}{D'''} + \frac{d'}{D'} + \frac{d''}{D''}} = \frac{f_0 - f}{R''' + R' + R''} \quad \dots\dots\dots (9)$$

the various  $d/D$  terms being the resistances to evaporation of the interfacial film ( $R'''$ ), the oil layer ( $R'$ ) and the stagnant air layer ( $R''$ ).

Equation (9) may be rewritten as

$$\frac{1}{Q}(f_0 - f) = R''' + R' + R'' \quad \dots\dots\dots (10)$$

where  $\frac{1}{Q}(f_0 - f)$  is the total evaporation resistance. Often it is more convenient to express the evaporation resistance in relative units, namely, compared with that of the free water surface of the same area, and at the same temperature and humidity. The resistance of a free water surface is due only to the stagnant air layer (Langmuir, 2; Rideal, 1). Since in this case  $R''' = 0 = R'$ ,

$$R'' = \frac{1}{Q_0}(f_0 - f) \quad \dots\dots\dots (11)$$

where  $Q_0$  is the rate of evaporation from the free water surface. Hence

$$\frac{Q_0}{Q} = \frac{R''' + R' + R''}{R''} = \frac{R'''}{R''} + \frac{R'}{R''} + 1 \quad \dots\dots\dots (12)$$

$Q_0/Q$  may be called the relative evaporation resistance.

Since for constant temperature  $R''$  is a constant,

$$Q_0/Q = \text{constant } (R''' + R' + R'') \quad \dots\dots\dots (13a)$$

or

$$\frac{Q_0}{Q} = \frac{R'''}{\text{const.}} + \frac{R'}{\text{const.}} + 1. \quad \dots\dots\dots (13b)$$

Approximations in this derivation are the assumptions (i) of absence of convection in the oil layer, (ii) that the stagnant air layer has a definite thickness and a constant fugacity gradient, and (iii) that the bulk layer of air above the stagnant layer is maintained at constant humidity.

Equations (12) and (13b) hold rigorously only when the free water surface is at the same temperature as the film covered surface. This may not always be

the case in comparative experiments because the temperature of the free water surface may be slightly lower than that of the film-covered one due to the difference in rate of evaporation. Equation (13a) is not subject to this restriction.

#### IV. EXPERIMENTAL

The paraffin oils used were commercial products (White Oil 890 of the Vacuum Oil Co., Nujol) as were the stand oil and the linseed oil used as spreaders. The eucalyptus residue consisted of the fractions boiling above 227° C. from the oil of *Eucalyptus amygdalina*. Polymerized oleic acid was prepared according to Heymann and Yoffe(5). The concentration of the spreader dissolved in the paraffin oil was 2 per cent.

Duplex films were spread on water in petri dishes of known area, filled to a constant level from the top. The calculated volume of oil was delivered from an "Agl" micrometer syringe. Spreading by this method may not be complete even after several days. Hence several drops of petroleum ether were added. The petroleum ether causes rapid spreading and evaporates very quickly, and uniform duplex films are obtained. The film thickness was calculated from the amount of oil placed on the surface and the area of the latter. The petri dishes were previously cleaned with sulphuric acid-dichromate cleaning mixture and washed with water free from surface-active material.

The evaporation tests were carried out over several hours, temperature and humidity being recorded. When not in use in evaporation experiments, the petri dishes with the films were kept under a glass cover to eliminate dust, and, in addition, the atmosphere was saturated with water vapour to prevent evaporation. In the absence of dust the rate of evaporation is usually constant over several weeks.

In many cases, in order to minimize the effects of different temperatures and humidities, the loss in weight per cm.<sup>2</sup> per hour through the film covered surface ( $Q$ ) was measured at the same time and under the same conditions as that for a free water surface ( $Q_0$ ). The rate of evaporation through the film was expressed relative to that of free water by the ratio  $Q_0/Q$ . Experiments carried out over considerable ranges of temperature (10-20° C.) and humidity (50-85 per cent.) gave values of  $Q_0/Q$  for each system which did not vary by more than 4-5 per cent. This method was used instead of the more rigorously defined one of Sebba and Briscoe in order to make experimental conditions conform to natural conditions. As the heavy paraffin oil itself suffered no measurable loss during the periods concerned,  $Q$  and  $Q_0$  were determined by weighing the petri dishes at regular intervals over a known period of time.

Evaporation tests were carried out in still air and in winds, at the same time and under the same general conditions of humidity and temperature. "Still air" experiments were performed in tall boxes to eliminate the effects of accidental draughts. The boxes were open at the top and contained duplicate film-covered surfaces and a free water control with a distance of at least 20 cm. between the dishes. If this distance was maintained the results were well reproducible and did not change on further increase of the distance.

Further "still air" experiments were performed in a large desiccator over 96 per cent. sulphuric acid,\* the film being spread as before. In order to keep the partial pressure of the sulphuric acid-water mixture constant, a large amount of 96 per cent. sulphuric acid and small dishes containing only little water were used. The partial pressure ( $p$ ) of water above sulphuric acid-water mixtures of 90-96 per cent. sulphuric acid is very small compared with the vapour pressure of pure water ( $p_0$ ), and varies little in that range ( $p < 0.14$  mm. Hg), so that  $p_0 - p \approx f_0 - f$  remains practically constant ( $p_0 = 15.5$  mm. Hg at  $18^\circ \text{C.}$ ) even if the water content of the sulphuric acid changes slightly during the experiment. Actually the rate of evaporation did not decrease with time.

Further evaporation tests were carried out in desiccators over 96 per cent. sulphuric acid and evacuated to 16 mm. Hg, which is a little above the pressure at which boiling occurs at room temperature. Air-free water was used in order to prevent the formation of air bubbles under the film.

Wind experiments were performed in a wind tunnel giving known wind speeds between 1 and 8 miles per hour†. The flow of air in this tunnel was non-turbulent. The rate of evaporation of water from a free surface was constant throughout the tunnel to  $\frac{1}{2}$  per cent. However, a reproducibility of only 4-5 per cent. could be obtained with film-covered surfaces. This is probably due to disturbance of the duplex film by the wind. Reference to Figure 2 shows that the rate of evaporation increased less than proportionately with the wind speed.

Preliminary wind experiments were performed with a blast of air of unknown speed from an electric fan.

## V. RESULTS AND DISCUSSION

The ratio of the rate of evaporation into a wind ( $Q^W$ ) to the rate of evaporation into still air ( $Q^S$ ) is shown for various thicknesses of the duplex film and wind velocities in Table 1,  $Q^W$  and  $Q^S$  being measured under identical conditions of temperature and humidity and expressed as g. cm.<sup>-2</sup> hr.<sup>-1</sup>. Stand oil and eucalyptus residue have been used as spreaders.

It is seen that the ratio  $Q^W/Q^S$  approaches unity at the highest measured wind speed at a duplex film thickness of  $10 \mu$ . Thus above a thickness of  $10 \mu$  a wind does not increase the rate of evaporation. If a fan is used  $Q^W/Q^S$  becomes unity only at 30-40  $\mu$  thickness; however, it is possible that, as a consequence of the turbulent nature of the air blast from the fan, the surface is agitated and the film disturbed, and the effective thickness of the duplex film may be less than the calculated value.

\* 96 per cent. sulphuric acid was chosen in preference to a salt hydrate because it was intended to have a gas phase of very low humidity.

† We are indebted to Professor J. S. Turner, Department of Botany, University of Melbourne, for the use of the wind tunnel.

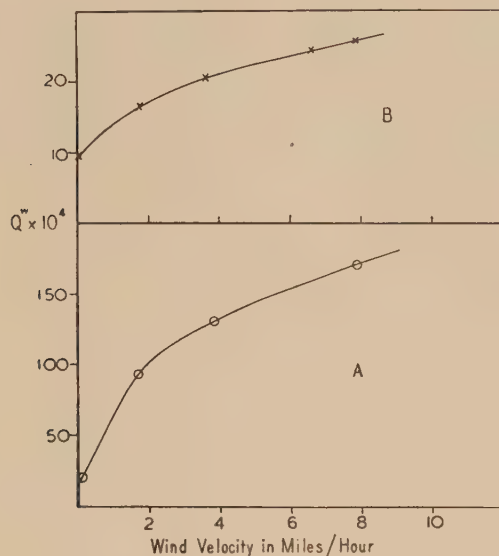


Fig. 2A.—Rate of evaporation ( $Q^w$ ) in  $\text{g. cm.}^{-2} \text{ hr.}^{-1}$  from a free water surface in relation to the wind velocity in miles  $\text{hr.}^{-1}$  (temperature,  $12\text{--}13^\circ \text{C.}$ ; humidity, 80–84 per cent.)

Fig. 2B.—Rate of evaporation ( $Q^w$ ) in  $\text{g. cm.}^{-2} \text{ hr.}^{-1}$  from a water surface covered by a duplex film,  $1 \mu$  thick (2 per cent. stand oil in paraffin oil), in relation to the wind velocity in miles  $\text{hr.}^{-1}$  (temperature  $11\text{--}13^\circ \text{C.}$ ; humidity, 65–72 per cent.).

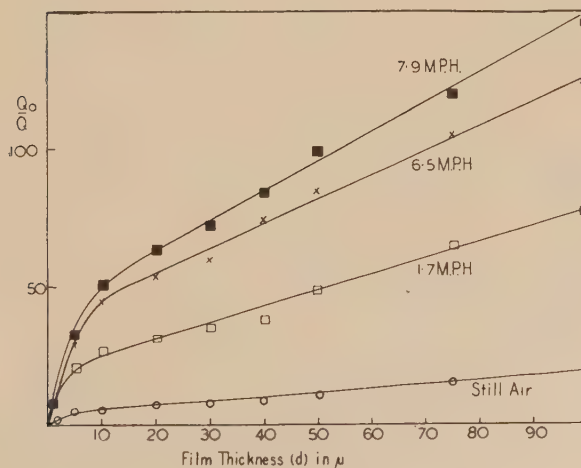


Fig. 3.—Plot of the relative evaporation resistance ( $Q_0/Q$ ) against the thickness ( $d$ ) of the duplex film for various wind velocities in miles  $\text{hr.}^{-1}$  (m.p.h.) and also for still air (2 per cent. eucalyptus residue in paraffin oil).



A wind increases the rate of evaporation because the stagnant layer of air is swept away and thus  $R''$  is reduced. Thus above a certain thickness it makes practically no difference to the rate of evaporation if the stagnant air layer is swept away by the wind. Above a duplex film thickness of about  $10\mu$  the resistance of the stagnant air layer ( $R''$ ) is thus small compared with that of the duplex film ( $R''' + R'$ ).

TABLE 1

RATIO OF THE RATE OF EVAPORATION INTO A WIND TO THE RATE OF EVAPORATION INTO "STILL AIR" ( $Q^w/Q^s$ ) UNDER IDENTICAL CONDITIONS OF TEMPERATURE AND HUMIDITY AT VARIOUS THICKNESSES OF THE DUPLEX FILM (DUPLICATES REPRODUCIBLE TO 4-5 PER CENT.)

Films	Wind Velocity (miles per hour)	Thickness of the Duplex Film ( $\mu$ )									
		0	1	5	10	20	30	40	50	75	100
2 per cent. stand oil in paraffin oil	1.7	4.55	1.74	1.01	0.99	0.99	0.96	0.98	1.04	0.96	0.99
	6.5	8.39	2.59	1.23	0.98	0.95	1.02	1.04	0.97	1.02	0.96
	7.9	9.10	2.90	1.29	1.01	1.03	1.01	0.97	1.04	0.99	0.98
	Fan	20.1	—	2.26	1.89	1.38	1.21	1.02	0.99	1.03	1.02
2 per cent. eucalyptus residue in paraffin oil	1.7	4.69	1.66	0.99	0.96	1.00	1.02	0.97	1.00	1.01	1.00
	6.5	8.48	2.19	1.12	0.95	1.00	0.98	0.99	1.02	0.98	1.00
	7.9	9.19	2.94	1.17	0.99	0.98	1.01	1.02	1.01	1.03	1.04
	Fan	20.3	—	2.31	1.67	1.25	1.10	1.01	1.02	1.01	1.04

Hence, considering equation (12),

$$\frac{Q_0}{Q} = \frac{R''' + R' + R''}{R''}, \dots\dots\dots (12)$$

in a wind,  $R''$  is reduced compared with conditions in still air, thus  $\frac{Q_0^w}{Q^w} > \frac{Q_0^s}{Q^s}$ ;

moreover  $\frac{Q_0^w}{Q^w}$  increases with increasing wind velocity. This is shown in Figure 3.

The greater  $Q_0/Q$ , the more efficient is a duplex film in reducing the rate of evaporation. Hence duplex films are more efficient in reducing evaporation in winds than in still air, and the efficiency increases with increasing wind velocity, provided the wind does not disturb the film. This was shown to be

the case with all spreaders under investigation. An analysis based on equation (12) also shows that a duplex film is more efficient in reducing evaporation into a vacuum than under atmospheric conditions, because  $R''$  is smaller in a vacuum than under atmospheric conditions. This can be seen on comparing the values for  $Q_0/Q$  in Figures 5 and 6. Rideal(1) obtained similar results for monolayers.

In several series of experiments  $Q_0/Q$  was determined at various film thicknesses ( $d$ ). For an analysis of the results, equations (12) and (13a) are rewritten

$$\frac{Q_0}{Q} = \frac{R'''}{R''} + 1 + \frac{d'/D'}{R''} \approx \frac{R'''}{R''} + 1 + K''d \quad \dots\dots\dots (14a)$$

and

$$\frac{Q_0}{Q} = K' \left( R''' + R'' + \frac{d'}{D'} \right) \approx K'(R''' + R'') + K''d \quad \dots\dots\dots (14b)$$

because  $d \approx d'$ , since  $d = d' + d''$  and  $d' \gg d''$ ;  $K'$  and  $K''$  are constants. The resistance of the stagnant layer of air above the film ( $R''$ ) is practically constant and independent of the film thickness.\*

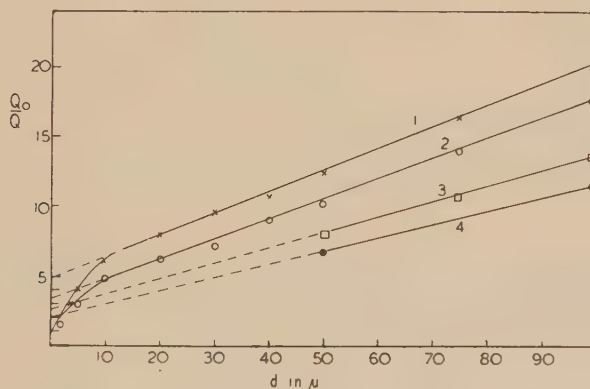


Fig. 4.—Plot of the relative evaporation resistance ( $Q_0/Q$ ) against the thickness ( $d$ ) of the duplex film in "still air" experiments at atmospheric pressure and humidities of 65–75 per cent. for :

1. 2 per cent. eucalyptus residue in paraffin oil.
2. 2 per cent. stand oil in paraffin oil.
3. 2 per cent. polymerized oleic acid in paraffin oil.
4. 2 per cent. linseed oil in paraffin oil.

The resistance of the interfacial film ( $R'''$ ) may be constant or may increase with increasing film thickness, because more spreader is available for adsorption

\* There may be a variation of  $R''$  with  $d$ , particularly in experiments carried out *in vacuo*, because the humidity will be greater above a thin duplex film than above a thick one. In experiments at atmospheric pressure and room temperature this variation, as can be shown easily, is not likely to be greater than 2–3 per cent.

from thicker duplex films than from thinner ones, the concentration of spreader in the oil being constant.\*

If  $R'''$  is independent of  $d$ , we obtain

$$Q_0/Q = K''' + K''d, \dots\dots\dots (15)$$

where  $K'''$  is another constant. Equation (15) is also obtained if  $R'''$  increases proportionally to  $d$ .

In all experiments  $Q_0/Q$  varies proportionally with  $d$  for film thicknesses over  $10\ \mu$ . This is shown by experiments in still air, keeping the evaporation dishes in open boxes in one series (Fig. 4), in a desiccator at atmospheric pressure over 96 per cent. sulphuric acid in another (Fig. 5), and in an evacuated desiccator over 96 per cent. sulphuric acid in a third series (Fig. 6). However, below  $d=10\ \mu$  the curves fall off.† This may be due to either (a) inhomogeneity

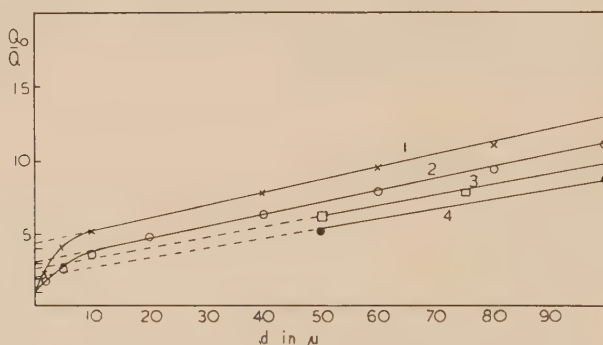


Fig. 5.—Plot of the relative evaporation resistance ( $Q_0/Q$ ) against the thickness ( $d$ ) of the duplex film in a desiccator at atmospheric pressure over 96 per cent. sulphuric acid for :

1. 2 per cent. eucalyptus residue in paraffin oil.
2. 2 per cent. stand oil in paraffin oil.
3. 2 per cent. polymerized oleic acid in paraffin oil.
4. 2 per cent. linseed oil in paraffin oil.

of the duplex films below  $d=10\ \mu$ , or (b) the resistance of the interfacial layer of spreader,  $R'''$  (*vide infra*), falling off very strongly at low thicknesses of the duplex film, because not enough spreader may be available for adsorption. In order to test these two possibilities the variation of  $Q_0/Q$  with  $d$  was investigated in circumstances in which the amount of spreader in the system is maintained constant. This was achieved by spreading pure paraffin oil to the required

\* An estimate of the amount of spreader adsorbed in the interface from the interfacial tension determined by Heymann and Yoffe(5) involves uncertainties. The interfacial tensions in systems containing polymerized spreaders may be in error because of the formation of a rigid skin in the interface.

† This is found also in experiments conducted in a wind tunnel (Fig. 3).

thickness on a basic film of  $1\ \mu$  thickness of the usual (2 per cent.) paraffin oil-spreader mixture (Fig. 7). However, in this case also, the curves fall off below  $d=10\ \mu$ . Thus the falling off of the curves in Figures 4-6 is more likely to be due to an inhomogeneity of the duplex films below  $10\ \mu$  thickness.

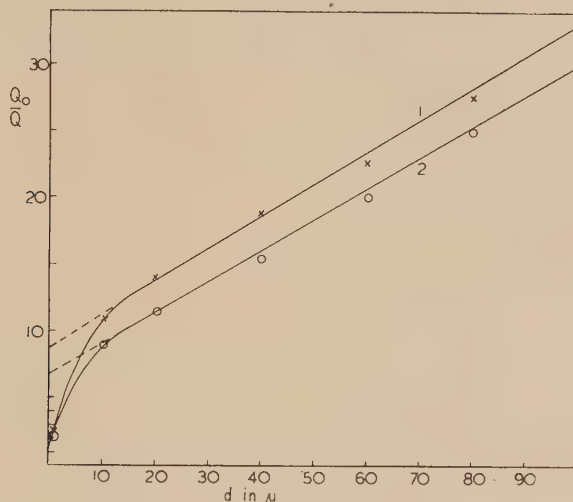


Fig. 6.—Plot of the relative evaporation resistance ( $Q_0/Q$ ) against the thickness ( $d$ ) of the duplex film in a vacuum desiccator (16 mm. Hg) over 96 per cent. sulphuric acid for :

1. 2 per cent. eucalyptus residue in paraffin oil.
2. 2 per cent. stand oil in paraffin oil.

Evidence in support of the inhomogeneity of thin duplex films is given by interference colours. The colours are not uniform, usually exhibiting red and green streaks and patches in random distribution. In some cases these duplex films, although apparently completely spread, exhibit Newton's rings when viewed by monochromatic light. In such cases, the duplex film is probably present as a very extended lens covering almost the whole surface, the film thickness being less towards the edges than in the centre.

Langmuir and Schaefer(4) have shown that with duplex films of "smoked" lubricating oils,  $R'''$  is negligible compared with  $R'$ . This is understandable in view of the fact that the amount of spreader present in these products is likely to be very small. In our duplex films, containing 2 per cent. of stand oil, linseed oil, eucalyptus residue, or polymerized oleic acid as spreaders, there is good evidence that the resistance due to the interfacial layer of spreader ( $R'''$ ) is quite appreciable compared with that of the layer of paraffin oil containing some dissolved spreader ( $R'$ ). This evidence emerges from the following observations :

(i) At all film thicknesses the value of  $Q_0/Q$  depends on the nature of the spreader. This is shown most clearly in Figures 4, 5, and 6. Reference to Figure 4 shows that, in experiments carried out under "still air" conditions, at a

duplex film thickness of  $50\ \mu$ , for instance,  $Q_0/Q$  is  $6.8$  for a duplex film containing linseed oil as a spreader,  $8.0$  for polymerized oleic acid,  $10.3$  for stand oil, and  $12.3$  for eucalyptus residue as spreaders. In all these duplex films, 2 per cent. solutions of the spreader and the same sample of paraffin oil were used. As the diffusion constant of water in paraffin oil (equations (14a) and (14b)) is unlikely to be altered appreciably by addition of 2 per cent. of spreader, the difference in the value of  $Q_0/Q$  between the duplex films containing these spreaders is most likely to be due to a difference in the resistance ( $R'''$ ) of the interfacial films of the various spreaders.\*

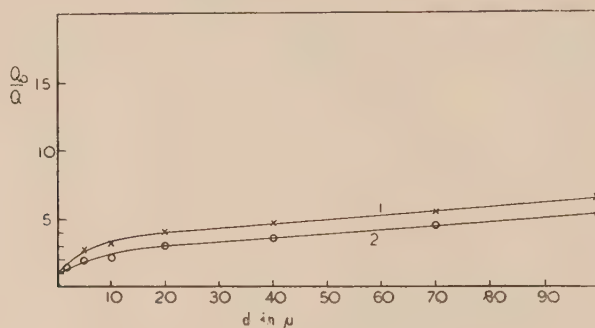


Fig. 7.—Plot of the relative evaporation resistance ( $Q_0/Q$ ) against the thickness ( $d$ ) of the duplex film in "still air" experiments at atmospheric pressure and a humidity of 60–65 per cent. In these experiments pure paraffin oil is spread to the required thickness on a basic film of  $1\ \mu$  thickness of the usual (2 per cent.) spreader-paraffin oil mixture.

Spreaders :

1. Eucalyptus residue.
2. Stand oil.

(ii) According to equation (14a) the intercepts ( $i$ ) on the  $Q_0/Q$  axes of the linear parts of the curves in Figures 4-6 are equal to  $\frac{R'''}{R''} + 1$ . The fact that the intercepts are greater than unity† shows that  $R'''$  must be appreciable compared with  $R'$  and  $R''$  (equation (13b)). The value of ( $i-1$ ) thus indicates how many times  $R'''$  is greater than  $R''$ . Because of the smallness of  $R''$  in a vacuum the value of  $i$  is much greater for the experiment in a vacuum desiccator than in the "still air" experiment. Table 2 shows values of ( $i-1$ ) which may

\* The dependence of  $Q_0/Q$  on the nature of the spreader might be explained on the assumption that the inhomogeneities, referred to above, may exist also with films that are thicker than  $10\ \mu$  and that they vary with the nature of the spreader. It would be difficult, however, to understand why, with strongly inhomogeneous films,  $Q_0/Q$  should vary linearly with the film thickness.

† The value of the intercept may be slightly smaller than unity if  $R'''$  is inappreciable, because the temperature of the free surface, due to more rapid evaporation, may be slightly lower than that of the film covered surface.



be regarded as relative measures of the evaporation resistance of the interfacial films of the various spreaders. Caution must be exercised, however, since it is assumed in this reasoning that  $R'''$  is independent of  $d$  above  $10\ \mu$  thickness.

TABLE 2  
VALUES OF  $(i-1)$  FOR VARIOUS SPREADERS

Spreader	"Still Air"	Desiccator over 96 per cent. $\text{H}_2\text{SO}_4$ (Atmospheric Pressure)	Desiccator over 96 per cent. $\text{H}_2\text{SO}_4$ (Vacuum)
Eucalyptus residue .. ..	3.9	3.2	8.0
Stand oil .. ..	2.4	2.0	5.7
Linseed oil .. ..	1.5	1.8	
Polymerized oleic acid ..	1.2	1.0	

These figures indicate that  $R'''$  is also quite appreciable compared with  $R'$ . Reference to Figure 4 shows that, for instance, for duplex films containing 2 per cent. stand oil,  $R'$  becomes greater than  $R'''$  only at a film thickness greater than 20 to  $30\ \mu$ .

(iii) Heymann and Yoffe(5) found that with duplex films containing eucalyptus residue the paraffin oil, after several months, often collected to lenses whilst a visible semi-solid skin remained. After this, there was still appreciable reduction of evaporation, obviously due to the remaining film. The film, at this stage, must have been multi-molecular and probably consisted of a three-dimensional network enmeshing some paraffin oil.

(iv) In the experiments previously referred to, in which pure paraffin oil was spread to the required thickness on a basic film of the usual 2 per cent. paraffin oil-spreader mixture of  $1\ \mu$  thickness (Fig. 7),  $Q_0/Q$  increases much less with increasing film thickness than in the experiments carried out in the usual way. This is probably the result of  $R'$  increasing with  $d$  as usually, whilst  $R'''$  decreases with  $d$ , because, when  $d$  is great, the amount of interfacial adsorption of spreader is less than when  $d$  is small, because the bulk concentration of spreader becomes smaller when more paraffin oil is spread on the basic paraffin oil-spreader film. Thus the results of these experiments also, are best explained on the assumption of an appreciable contribution by the interfacial layer of spreader to the evaporation resistance of the duplex film.

## VI. REFERENCES

- (1) RIDEAL, E. K.—*J. Phys. Chem.* **29**: 1585 (1925)
- (2) LANGMUIR, I., and LANGMUIR, D. B.—*Ibid.* **31**: 1719 (1927).
- (3) SEBBA, F., and BRISCOE, H. V. A.—*J. Chem. Soc.* **1940**: 106 (1940).

- (4) LANGMUIR, I., and SCHAEFER, V. J.—*J. Franklin Inst.* **235**: 119 (1943).
- (5) HEYMANN, E., and YOFFE, A.—*Nature* **146**: 265 (1940); *Trans. Faraday Soc.* **38**: 408 (1942); *Ibid.* **39**: 217 (1943).
- (6) HARKINS, W. D.—*J. Chem. Phys.* **9**: 552 (1941).
- (7) HEYMANN, E., and YOFFE, A.—*J. Phys. Chem.* **49**: 239 (1945).
- (8) MERCER, E. H.—*Proc. Phys. Soc. Lond.* **51**: 561 (1939).
- (9) MILLER, N. F.—*J. Phys. Chem.* **45**: 1025 (1941).
- (10) LANGMUIR, I., and SCHAEFER, V. J.—*J. Amer. Chem. Soc.* **59**: 2403 (1937).
- (11) BARRER, R. M.—“Diffusion in and through Solids”, p. 4. (Cambridge, 1941.)

# EQUILIBRIA IN FURFURAL-WATER SYSTEMS UNDER INCREASED PRESSURE AND THE INFLUENCE OF ADDED SALTS UPON THE MUTUAL SOLUBILITIES OF FURFURAL AND WATER

By R. G. CURTIS\* and H. H. HATT\*

[Manuscript received February 10, 1948]

## Summary

Liquid-vapour equilibrium data are reported for the furfural-water system at pressures of 6·17, 7·85, and 9·83 kg. per cm.<sup>2</sup>† Under these pressures the system forms a homoazeotropic mixture differing only slightly in composition from the heteroazeotropic mixture formed at normal pressures. The data are checked by the calculation of activity coefficients. Abnormal activity coefficients are found for water in mixtures of high furfural content.

The equation

$$\log_{10} P = -\frac{2209}{T} + 7.959$$

fits published data and that now reported for the vapour pressure of furfural over the range of pressures 100 to 6,000 mm. Hg.

Changes in the mutual solubilities of furfural and water caused by the addition of metallic sulphates are determined. All the sulphates examined raised the critical solution temperature, sodium sulphate being most effective. With magnesium sulphate there was a region of discontinuity, which may possibly be attributed to a change of phase or of hydration of magnesium sulphate.

The velocities of reaction (decomposition or polymerization) of furfural in aqueous and in aqueous acidic medium at 180° C. are reported. The data are used to estimate the amount of furfural lost by such reactions in a commercial process for the production of furfural. It is considered that the use of metallic sulphates in the commercial production of furfural should not influence the yield appreciably.

## INTRODUCTION

In the commercial production of furfural, plant material is digested with dilute mineral acid under pressure and the furfural formed is removed continuously in a current of steam. Conversion of the pentosans present to furfural is far from quantitative, the furfural obtained corresponding only to 50–60 per cent. of the pentosans used. Probably losses occur at each stage in the conversion of pentosans to furfural through side reactions, but there is also loss through subsequent transformation of the furfural. The data presented in this

\* Division of Industrial Chemistry, C.S.I.R.

† The equivalent English values for these three pressures are 87·7, 111·7, and 139·7 lb. per sq. in.

paper allow more exact estimates of the furfural lost by subsequent reaction and provide some basis from which to calculate the improvements in yield to be expected from changes in the method of digesting plant material.

The experimental data fall into three parts. First, information is given concerning liquid-vapour equilibria for furfural-water systems at pressures above atmospheric. Hitherto, data for this partially miscible system have been reported only at atmospheric pressure when furfural and water form a heteroazeotrope boiling at  $97.9^{\circ}\text{C}$ . and containing 35 per cent. by weight of

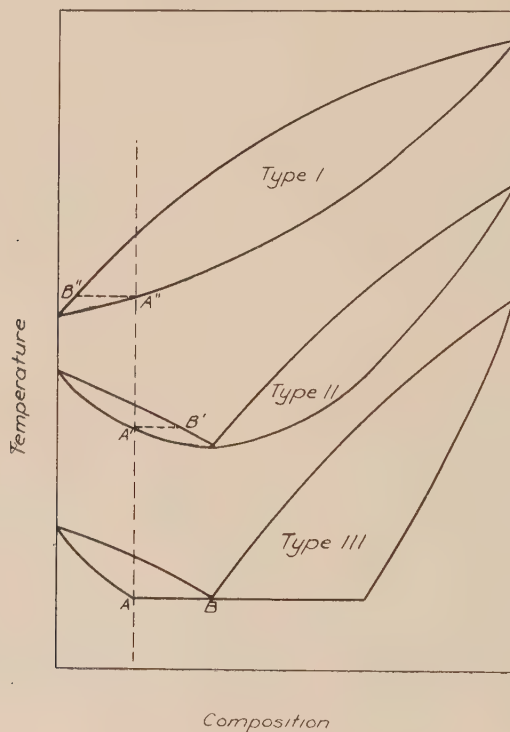


Fig. 1

furfural(1). The critical solution temperature, which has been carefully determined by Evans and Aylesworth(2), lies at  $120.9^{\circ}\text{C}$ . In the commercial production of furfural the pressures used vary from 7.0 to 10.5 kg. per  $\text{cm}^2$ , and from the shape of the distillation isobars at atmospheric pressure it was clear that at these higher pressures all furfural-water mixtures would be completely miscible at their boiling points and the distillation isobars would presumably then belong to Type II of Figure 1, showing a minimum boiling point. However, departure from Raoult's law decreases as the pressure on a system increases and the manifestation of azeotropy becomes increasingly difficult as the disparity between the boiling points of the components increases. At normal pressure the difference

between the boiling points of azeotrope and water is  $2.1^{\circ}\text{C.}$  while between furfural and water it is  $62^{\circ}\text{C.}$  and increases with increasing pressure, so that at high pressures the system might show distillation isobars of Type I.

Whether at these higher pressures the distillation isobars conformed to Type I or Type II, it seemed worth while to examine the possibility of causing reversion to the heteroazeotropic Type III by the addition of a third component. In the phenol-water system, for example, the addition of certain metallic salts causes a marked extension of the region of immiscibility and a great increase in the critical solution temperature. Similar changes with the furfural-water system would maintain distillation isobars of Type III at high pressures with an accompanying enrichment in furfural of the vapour in equilibrium with liquid of a given furfural-water composition. This would produce an accelerated elimination of furfural upon distillation, appreciable in changing from Type II to Type III, but much greater and of considerable value if the change were from Type I to Type III (compare in Figure 1 points  $A$ ,  $A'$ , and  $A''$  on the boiling point curves and their conjugate points  $B$ ,  $B'$ , and  $B''$  on the vapour composition curve). The second part therefore describes the changes in the mutual solubilities of furfural and water caused by the addition of certain metallic sulphates.

Lastly, we have examined the rate of decomposition of furfural at  $180^{\circ}\text{C.}$  in the presence of water and in the presence of dilute sulphuric acid. Little has been published on the behaviour of furfural in dilute acid at high temperatures. Dunlop and Peters(3) found dry furfural to be relatively stable in absence of air up to  $230^{\circ}\text{C.}$  and Lazarev(4) examined its rate of destruction by mineral acids, but only at  $100^{\circ}\text{C.}$  When the velocity of decomposition of furfural in aqueous solution is known, it should be possible to use these data together with those on liquid vapour equilibria to calculate the percentage of furfural lost in commercial production through such decomposition or polymerization reactions.

## PART I.—LIQUID-VAPOUR EQUILIBRIA IN FURFURAL-WATER SYSTEMS

### (a) *Apparatus and Method*

An equilibrium still was used of the type described by Griswold, Andres, and Klein(5), but with some modifications made necessary by the need to condense the distillate above the critical solution temperature in order to withdraw samples of a homogeneous liquid. The apparatus was constructed of brass and bronze and is shown diagrammatically in Figure 2. The modifications introduced were these.

A metallic bellows  $E$  and restraining spring were inserted at the top of the still  $B$ , and by expansion and contraction actuated a micro-switch which controlled the current to the heater leg  $A$ . This provided the main control of temperature and of the rate of distillation and could be adjusted by varying the height of the micro-switch by means of a fine screw. The bellows was placed within the still to avoid unnecessary condensation. A thin copper cylinder  $C$  placed in the upper portion of the still prevented any enriched vapour, produced by partial condensation at the outer wall, gaining access to the condenser.



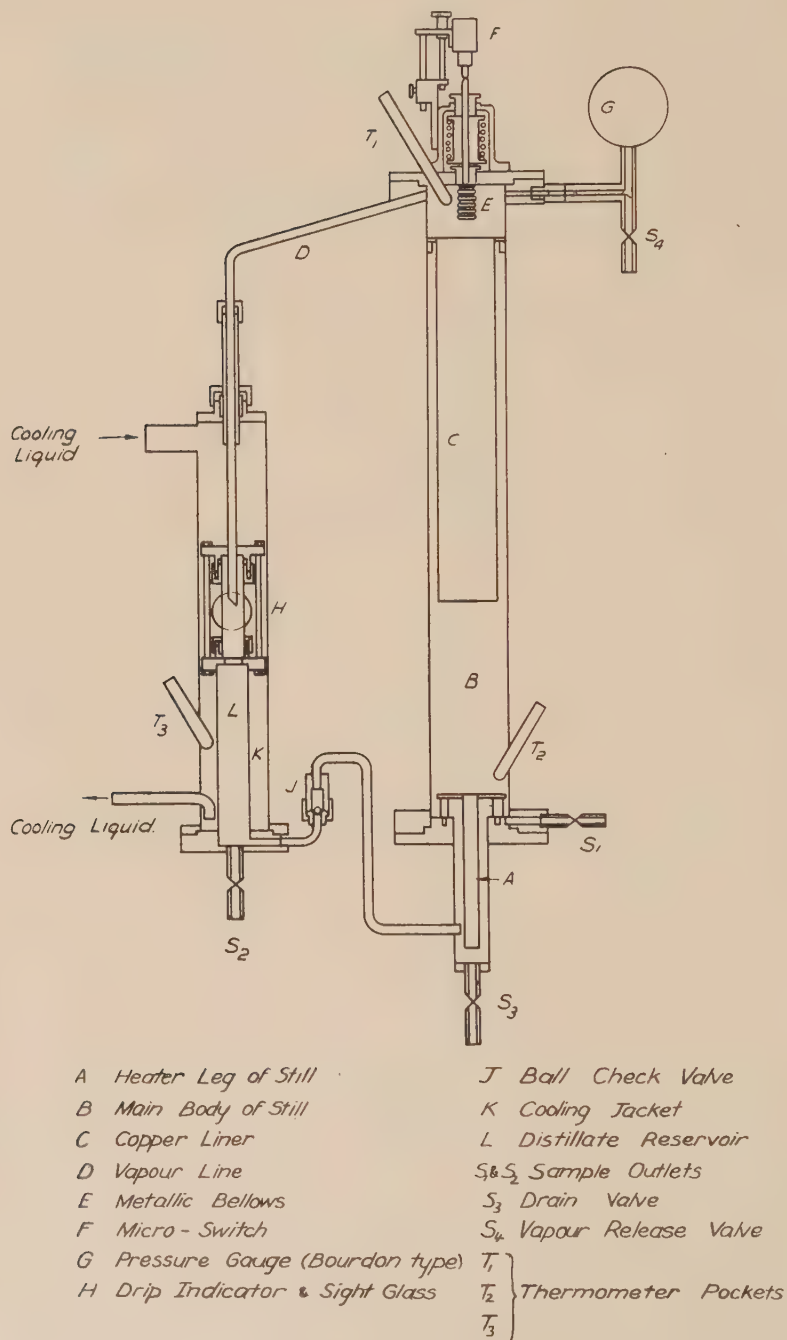


Fig. 2

The upper and lower plates of *B* were not welded, but bolted on so that the heater leg *A* could be freed when necessary from deposits of resinified furfural. Separate heating elements were wound for the upper and lower portions of the still and were controlled through Variac transformers. Four copper-constantan thermocouples, equally spaced, were welded to the wall of the still and accurate thermometers, graduated in fifths of degrees, were inserted in the thermometer pockets. The whole still was heavily lagged with asbestos tape.

In the check valve *J*, a  $\frac{3}{16}$  in. steel ball was too heavy, preventing an even flow of distillate to the still, and was replaced by a glass ball of the same dimensions.

A single brass condenser *K* was used and between the vapour tube *D* and the reservoir *L* a piece of pyrex pressure glass was inserted, enabling the rate of distillation to be observed. The vapour tube was cut obliquely and served as a drip indicator. Seatings for the glass tube were of neoprene reinforced with fabric and lasted for several weeks. The condenser *K* was kept at 130°C. by circulating a mixture of liquid and vapour generated in a modified Cottrell boiler (not shown in Figure 2). The liquid used was a petroleum ether fraction of suitable boiling point. The condenser *K* was also lagged.

In operation, the pressure gauge *G*, when hot, gave readings from 0.15 to 0.30 kg. higher than those obtained by calibration when cold. Therefore, before and after a new set of experiments, water was distilled in the apparatus, and from the temperature of the still the true pressure was calculated. A run was commenced by evacuating the apparatus to 0.5 mm. and then introducing 600 ml. of a suitable furfural-water mixture through valve *S*<sub>3</sub>, avoiding entrance of air and once more quickly evacuating to 0.5 mm. The apparatus was heated, the condenser liquid circulated and inputs to the heating elements on *B* adjusted so that the thermocouples read 4°C. higher than thermometers *T*<sub>1</sub> and *T*<sub>2</sub>, which themselves agreed to within less than 0.5°C. The temperature at *T*<sub>1</sub> was accepted as the distillation temperature.

The rate of distillation depended on the difference of temperature between the still and the condenser. The minimum practicable rate was from 60 to 75 drops per minute, being determined by the need to replace the distillate in the reservoir completely several times before taking a sample. Several samples were taken in a run at intervals of 1-1½ hours, and because of resinification a run could not last more than a few hours. Samples of residual liquid and distillate were withdrawn simultaneously through valves *S*<sub>1</sub> and *S*<sub>2</sub> into 20 ml. bombs of the type described by Griswold, Andres, and Klein, and after cooling quickly and weighing, these were diluted with water, to an extent dependent upon the amount of furfural present, and the furfural estimated by titration with bromide and bromate. Gravimetric check analyses with thiobarbituric acid by the method of Mackney and Reynolds(6) were made occasionally.

The furfural used in these experiments was purified in the way described by Evans and Aylesworth(2) and was redistilled under reduced pressure immediately before use.

(b) *Results*

Observations covering the whole composition range have been made at absolute pressures of 6.17 and 7.85 kg. per cm.<sup>2</sup>. Observations have also been made at 9.83 kg. per cm.<sup>2</sup> pressure for compositions rich in water, and of more interest to us. Observations covering the whole composition range could not be made at this pressure, for mixtures rich in furfural boiled at high temperatures

TABLE 1A  
FURFURAL-WATER MIXTURES

BOILING POINTS AND EQUILIBRIUM LIQUID AND VAPOUR COMPOSITIONS AT 6.17 KG. PER CM.<sup>2</sup>

Boiling Point (° C.)	Liquid		Vapour	
	Furfural (%)	Furfural (mole fraction)	Furfural (%)	Furfural (mole fraction)
159.0	0.0	0.000	0.0	0.000
158.8	0.9	0.0017	3.6	0.0070
158.7	0.9	0.0017	4.5	0.0088
158.1	4.4	0.0086	14.8	0.0316
158.0	4.8	0.0094	17.0	0.0370
157.5	7.7	0.0154	25.3	0.0598
157.3	11.2	0.0231	27.6	0.0668
157.0	20.8	0.0470	31.9	0.0808
156.95	29.6	0.0731	34.2	0.0889
156.9	34.6	0.0903	34.6	0.0903
157.1	49.4	0.155	36.5	0.0974
157.3	70.8	0.313	38.0	0.103
158.6	79.6	0.423	39.6	0.110
159.4	83.1	0.480	39.5	0.109
162.4	87.9	0.577	41.0	0.115
174.0	92.1	0.687	52.4	0.171
176.0	93.2	0.720	56.6	0.197
178.8	92.2	0.689	58.2	0.207
180.0	93.9	0.743	60.6	0.224
189.6	95.2	0.788	66.5	0.271
201.0	96.2	0.827	74.2	0.350
204.0	96.3	0.830	78.5	0.406
239.8	100	1.000	100	1.000

and resinified rapidly. At 6.17 and 7.85 kg. pressure, mixtures rich in furfural, because of their higher boiling points, distilled too rapidly, and to reduce the rates a cooling liquid of higher boiling point was circulated. Hexanol was suitable for this purpose. The furfural-water system was also investigated at a

pressure of 4.55 kg. per cm.<sup>2</sup>, but the mixtures rich in water boiled at about 145° C. and distillation rates were too slow for reliable homogeneous samples to be obtained.

The data obtained are collected in Tables 1A, 1B, and 1C and shown graphically in Figures 3, 4, and 5. They show that up to a pressure of 10 kg.

TABLE 1B

## FURFURAL-WATER MIXTURES

BOILING POINTS AND EQUILIBRIUM LIQUID AND VAPOUR COMPOSITIONS AT 7.85 KG. PER CM.<sup>2</sup>

Boiling Point (° C.)	Liquid		Vapour	
	Furfural (%)	Furfural (mole fraction)	Furfural (%)	Furfural (mole fraction)
168.9	0.0	0.000	0.0	0.000
168.65	1.0	0.0019	4.1	0.0080
168.6	1.3	0.0025	4.6	0.0090
167.9	4.3	0.0084	15.1	0.0323
167.8	4.8	0.0094	15.8	0.0340
167.4	8.3	0.0167	22.2	0.0508
167.3	8.8	0.0178	24.9	0.0586
167.25	11.4	0.0236	25.7	0.0610
167.2	11.9	0.0247	25.6	0.0607
166.8	22.5	0.0517	32.6	0.0832
166.75	28.6	0.0699	34.2	0.0889
166.7	37.2	0.100	34.7	0.0907
166.75	40.0	0.111	35.1	0.0921
166.8	41.1	0.116	35.5	0.0936
167.0	54.3	0.182	35.8	0.0948
167.5	66.7	0.273	36.8	0.0985
167.6	69.0	0.295	37.9	0.103
168.8	78.0	0.400	39.5	0.109
169.2	80.1	0.430	40.0	0.111
173.0	86.1	0.538	45.0	0.133
174.2	86.6	0.548	46.8	0.142
183.2	92.4	0.695	51.8	0.168
186.4	93.0	0.714	55.1	0.187
194.4	94.2	0.754	62.4	0.237
198.0	95.8	0.811	67.2	0.278
210.0	96.7	0.846	74.9	0.359
213.4	96.7	0.846	75.5	0.366
215.0	96.8	0.851	77.1	0.387
224.0	97.6	0.884	85.4	0.524
253.8	100	1.000	100	1.000

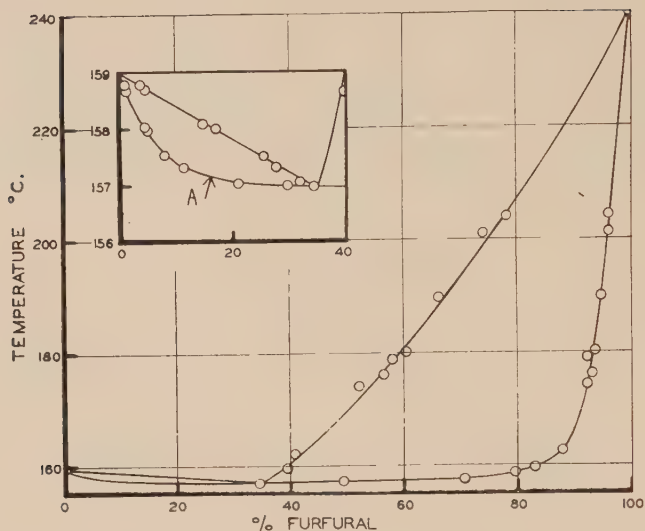
Fig. 3.—Liquid-vapour equilibrium curve at 6.17 kg./cm.<sup>2</sup>

TABLE IC

## FURFURAL-WATER MIXTURES

BOILING POINTS AND EQUILIBRIUM LIQUID AND VAPOUR COMPOSITIONS AT 9.83 KG. PER CM.<sup>2</sup>

Boiling Point (° C.)	Liquid		Vapour	
	Furfural (%)	Furfural (mole fraction)	Furfural (%)	Furfural (mole fraction)
178.3	0.0	0.000	0.0	0.000
178.2	0.8	0.0015	3.5	0.0068
178.1	1.0	0.0019	3.8	0.0074
177.8	2.7	0.0052	9.5	0.0193
177.7	3.0	0.0058	10.0	0.0204
177.6	3.6	0.0070	11.8	0.0245
177.5	4.5	0.0088	13.5	0.0284
177.2	7.3	0.0146	18.7	0.0414
176.8	10.8	0.0222	26.1	0.0622
176.7	12.8	0.0268	26.4	0.0631
176.5	18.1	0.0398	31.0	0.0778
176.4	20.4	0.0459	31.6	0.0798
176.3	27.8	0.0674	33.9	0.0878
176.2	33.8	0.0874	34.6	0.0903
176.3	37.5	0.101	35.2	0.0925
176.8	54.7	0.185	36.4	0.0970



the furfural-water system belongs to Type II of Figure 1 and exhibits no tendency to pass to Type I with further increase of pressure. In fact, except that the furfural-water system becomes completely miscible before the boiling-point curve is reached, the changes introduced by increased pressure are small. The

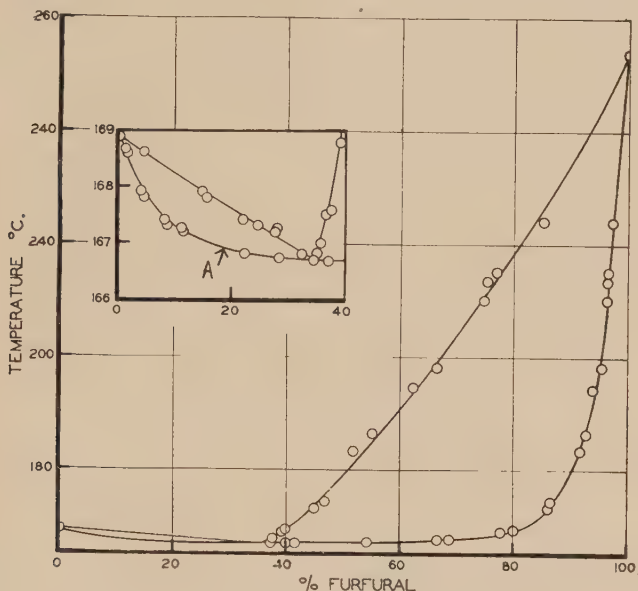


Fig. 4.—Liquid-vapour equilibrium curve at 7.85 kg./cm.<sup>2</sup>

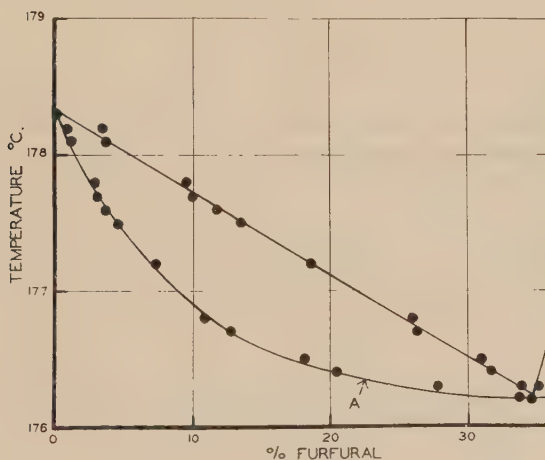


Fig. 5.—Liquid-vapour equilibrium curve at 9.83 kg./cm.<sup>2</sup>

difference between the boiling point of the azeotropic mixture and water is still the same as at normal pressures:  $2.1^{\circ}\text{C}$ . The compositions of the homoazeotropic mixtures formed at higher pressures differ from that of the heteroazeotrope at atmospheric pressure by an amount scarcely outside the

experimental error. Mains(1) gave the mole fraction of furfural in the hetero-azeotrope as 0.0918; for the homoazeotrope we obtained values of 0.0903 to 0.0910. The diminished departure from Raoult's law by systems under increased pressure is shown in the plot of vapour composition against liquid composition in Figure 6.

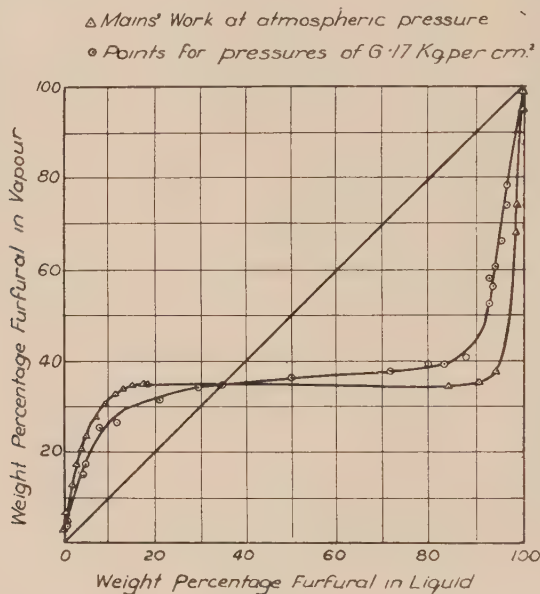


Fig. 6

### (c) Discussion

We have attempted to check the reliability of our experimental data and particularly the small variation in the composition of the azeotrope with pressure, by the use of certain empirical formulæ and by the calculation of activity coefficients. For this purpose the vapour pressures of furfural at various temperatures were needed. These were obtained from an empirical equation based on data reported by Evans and Aylesworth(2) and results obtained by us at higher pressures.

The equation

$$\log_{10} P = -\frac{2209}{T} + 7.959 \quad \dots \dots \dots (1)$$

was fitted by the method of least squares, and using Bessel's formula

$$r = \pm 0.6745 \sqrt{\frac{\sum d^2}{n-k}},$$

the probable error was 0.44° C. The observed and calculated boiling points are compared in Table 2. Evans and Aylesworth's observation at 69 mm. differed by more than four times the probable error from the value calculated from the equation obtained when this observation was included. It was therefore omitted when deriving equation (1).

TABLE 2†  
 BOILING POINTS OF FURFURAL

Pressure (mm. Hg)	Boiling Point (° C.)		Difference
	Observed	Calculated (Equation (1))	
8	39.9	39.9	0.0
69	92.3	87.8	+4.5
214	120.3	119.2	+1.1
310	131.6	130.8	+0.8
411	140.2	140.1	+0.1
625	154.4	154.7	-0.3
707	159.0	159.1	-0.1
744	160.9	161.1	-0.2
760	161.7	161.8	-0.1
812	163.8	164.3	-0.5
966	170.6	170.9	-0.3
3357*	223.8	225.1	-1.3
4551*	239.8	240.4	-0.6
5777*	253.8	253.3	+0.5

\* Data not asterisked are taken from Evans and Aylesworth.

† There has been no careful measurement of the low vapour pressures of furfural. Evans and Aylesworth's measurements, which are the best available, appear unreliable at low pressures. Their observation at 69 mm., which we reject, may be more reliable than that at 8 mm., for our linear relationship (equation (1)) may not hold over an extended range and for this may need replacement by an equation of the type

$$\log P = -\frac{A}{T} + B \log T + C.$$

Single observations of vapour pressures support this view. Desner(7) reports b.p.<sub>9</sub> 60° C.; Landrieu, Baylocq, and Johnson(8), b.p.<sub>13</sub> 60-61° C.; Walden(9), b.p.<sub>25</sub> 72° C. and Timmermanns(10), b.p.<sub>87</sub> 77° C. All are at variance with the b.p.<sub>8</sub> 39.9° C. of Evans and Aylesworth. But the data are not accurate or consistent enough to warrant an equation of the second order.

Recently, Stull(11) assigned vapour pressure values for furfural apparently based solely on Walden's single determination and the boiling point at atmospheric pressure. He appears to have overlooked or else discarded the most recent work: that of Evans and Aylesworth. Their measurements at near atmospheric pressure appear reliable and have been combined with our own for use in equation (1) at higher pressures. Stull's assigned values differ considerably from those of Evans and Aylesworth as the following comparison shows:

Pressure (mm. Hg)	B.P. (Stull)	B.P. (Eqn. (1))
1	18.5	4.4
5	42.6	31.1
10	54.8	44.3
20	67.8	58.7
40	82.1	74.4
60	91.5	84.3
100	103.4	97.6
200	121.8	117.3
400	141.8	139.2
760	161.8	161.8

Equation (1) gives a latent heat of evaporation of 105.3 in fair agreement with the observed value of 107.5 cal./g.

Roozeboom's empirical rule states that with increasing pressure the composition of the azeotropic mixture is enriched in that component having the steeper vapour pressure curve. In the present example it is water

$$\left(\frac{dp}{dt}\right)_{\text{H}_2\text{O}}^{100^\circ \text{C.}} = 27.1, \quad \left(\frac{dp}{dt}\right)_{\text{furfural}}^{100^\circ \text{C.}} = 4.0 \text{ mm.Hg/}^\circ \text{C.,}$$

and with this our findings agree. However, there are many exceptions to this rule and Prigogine(12), while giving it a thermodynamical basis, showed it must have very limited application. By thermodynamical methods Prigogine has developed the equation

$$\frac{N_2}{1-N_2} = \sqrt{\frac{\log P/P_{10}}{\log P/P_{20}}} \dots\dots\dots (2)$$

which should have wider application. Here  $P$  is the total pressure of the system,  $P_{10}$  and  $P_{20}$  are the pressures of the pure components at the boiling point of the azeotrope, and  $N_2$  is the mole fraction of furfural in the azeotropic mixture. Equation (2) fails to predict the composition of the furfural-water azeotrope correctly, giving  $N_2=0.162$  at 760 mm. and 0.146 at 5,777 mm. instead of the observed 0.092 and 0.091.

Carlson and Colburn(13) recommend the use of van Laar's equations (3a) and (3b) if the boiling point of the azeotrope at one pressure is known.

$$\log \gamma_1 = \frac{A}{\left(1 + \frac{AN_1}{BN_2}\right)^2} \dots\dots\dots (3a)$$

$$\log \gamma_2 = \frac{B}{\left(1 + \frac{BN_2}{AN_1}\right)^2} \dots\dots\dots (3b)$$

The method assumes that the van Laar constants  $A$  and  $B$  are independent of temperature. Using Mains' data at 760 mm. the values for the van Laar constants were found as  $A=0.448$  and  $B=1.644$ . Vapour pressures of furfural were obtained from equation (1) and with those of water were used to calculate the activity coefficients, the compositions of the azeotrope at increased pressures then being obtained from the equation

$$\frac{N_2}{1-N_2} = \frac{\log \gamma_1}{\log \gamma_2} \left( \sqrt{\frac{A}{\log \gamma_1}} - 1 \right)$$

The results of these calculations show fair agreement with our experimental work and are notable for predicting only a slight change in azeotropic composition with increasing pressure.

Pressure (kg. per cm. <sup>2</sup> )	..	1.03	6.17	7.85	9.83
B.P. of Azeotrope (° C.)	..	97.9	156.9	166.7	176.3
Mole Fraction of Furfural					
(Observed)	.. ..	0.0918	0.091	0.091	0.091
Mole Fraction of Furfural					
(Calculated)	.. ..	—	0.087	0.087	0.085

The activity coefficients of water and furfural in the liquid phase have been calculated from the simple equation

$$\gamma = \frac{y_1 P}{x_1 P_1}$$

where  $x_1$  and  $y_1$  are the mole fractions of the component in the liquid and vapour phases,  $P$  is the pressure on the system, and  $P_1$  the vapour pressure of the pure

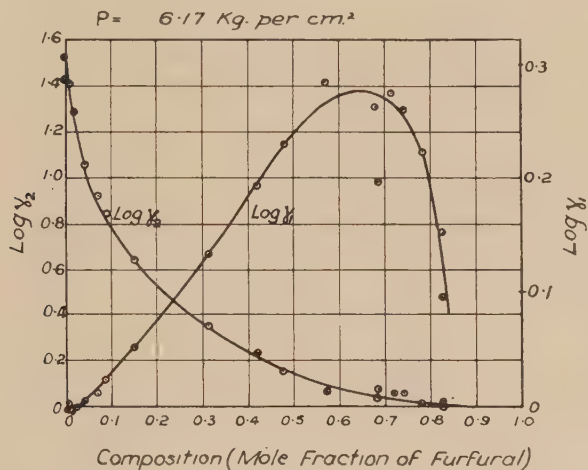


Fig. 7

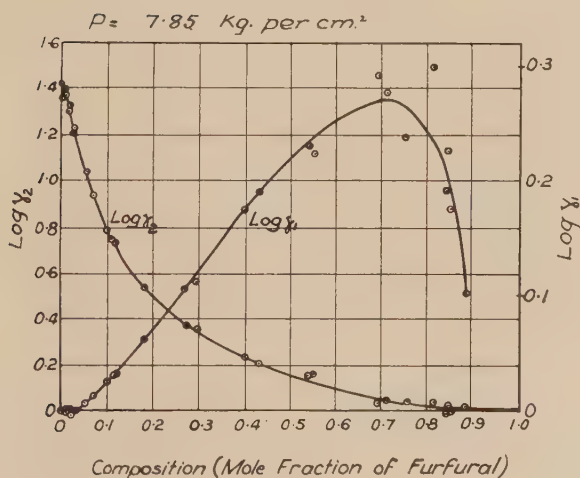


Fig. 8

component at that temperature, and are given in Table 3 for pressures of 6.17 and 7.85 kg. Following Carlson and Colburn, in order to reveal irregularities, the logarithms of these coefficients have been plotted against composition (Figs. 7 and 8). These figures make the irregularities in the data for mixtures of high furfural content very apparent. Elsewhere the data fit the requirements :



TABLE 3  
 ACTIVITY COEFFICIENTS

Pressure = Atmospheric (Data calculated from Work of Mains)			Pressure = 6.17 kg. per cm. <sup>2</sup>			Pressure = 7.85 kg. per cm. <sup>2</sup>		
Mole Fraction of Furfural	Activity of Water ( $\gamma_1$ )	Activity of Furfural ( $\gamma_2$ )	Temp.* (° C.)	Activity of Water ( $\gamma_1$ )	Activity of Furfural ( $\gamma_2$ )	Temp.* (° C.)	Activity of Water ( $\gamma_1$ )	Activity of Furfural ( $\gamma_2$ )
0.0015	0.998	71.05	159.0	1.000	—	168.9	1.000	—
0.0026	1.002	55.96	158.8	1.000	26.57	168.65	0.999	26.68
0.0049	0.996	50.78	158.7	1.000	33.49	168.6	1.000	22.86
0.0068	0.995	53.33	158.1	1.000	24.16	167.9	0.998	24.87
0.0099	1.002	44.60	158.0	0.998	25.95	167.8	1.000	23.45
0.0138	1.002	37.00	157.5	0.993	25.96	167.4	1.000	19.94
0.0209	1.004	29.90	157.3	0.998	19.44	167.3	0.995	21.63
0.0340	1.015	19.88	157.0	1.015	11.65	167.25	1.000	17.00
0.0477	1.030	14.17	156.95	1.036	8.250	167.2	1.002	16.18
0.0744	1.060	9.127	156.9	1.055	6.794	166.8	1.016	10.71
0.0906	1.077	7.586	157.1	1.121	4.247	166.75	1.031	8.476
0.169	1.177	4.072	157.3	1.363	2.212	166.7	1.065	6.053
0.336	1.476	2.043	158.6	1.558	1.687	166.75	1.075	5.530
0.464	1.825	1.505	159.4	1.696	1.441	166.8	1.078	5.371
0.585	2.102	1.287	162.4	1.919	1.166	167.0	1.158	3.448
0.633	2.624	1.099	174.0	1.831	1.076	167.5	1.281	2.358
0.774	3.986	0.941	176.0	1.891	1.124	167.6	1.312	2.276
0.916	4.505	1.154	178.8	1.575	1.151	168.8	1.487	1.721
0.925	3.442	1.035	180.0	1.815	1.121	169.2	1.546	1.614
0.964	0.795	1.026	189.6	1.667	1.014	173.0	1.698	1.402
0.984	0.873	0.974	201.0	1.426	0.957	174.2	1.669	1.424
0.997	0.942	1.017	204.0	1.247	1.035	183.2	1.948	1.062
1.000	—	1.000	239.8	—	1.000	186.4	1.889	1.064
						194.4	1.730	1.059
						198.0	1.974	1.061
						210.0	1.681	1.005
						213.4	1.554	0.951
						215.0	1.505	0.967
						224.0	1.263	1.043
						253.8	—	1.000

\* Temperature serves to connect with the data of Table 1.

the activity coefficient of water approaches unity regularly and that of furfural attains a maximum as the mole fraction of furfural approaches zero;  $\log \gamma_2$  is considerably less than  $\log \gamma_1$  at a mole fraction of 0.5.

The discrepancies above  $N_2=0.7$  are of two kinds; there is considerable scattering of readings and there is a marked maximum in  $\log \gamma_1$  without a

corresponding minimum for  $\log \gamma_2$ . It is certain that some of these discrepancies are due to experimental error, for these regions of the curves correspond to the mixtures of highest boiling point, where decomposition of furfural is rapid. Moreover, the change in boiling point with temperature is greatest in this region and therefore the chances of failing to attain equilibrium conditions greatest. At 7.85 kg. while the liquid composition changes from 93 to 96 per cent. furfural, the boiling point rises 25° C.

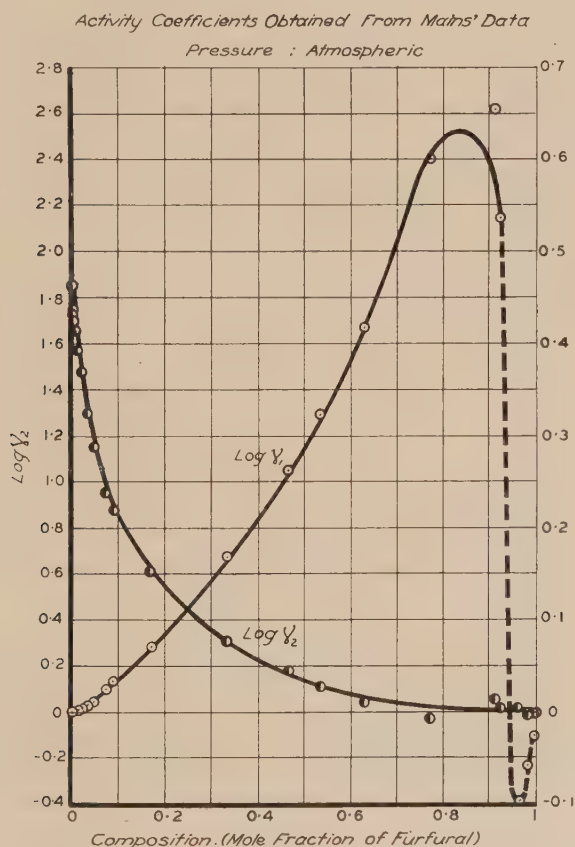


Fig. 9

That the maxima are real and not caused by an arbitrary location of the curve amidst the scattered points was shown by reading values of  $\log \gamma_2$  from smoothed curves (without minima) and calculating values of  $\gamma_1$  from the equation

$$\gamma_1 = \frac{\gamma_2 x_2 y_1 P_2}{x_1 y_2 P_1}$$

When the derived values of  $\log \gamma_1$  were plotted, the maxima were still obtained. Furthermore Mains' data for liquid-vapour equilibria at atmospheric pressure have been used to calculate activity coefficients, and the derived values of  $\log \gamma_1$

also show a maximum—in this case at a furfural molar fraction of about 0.8. For mixtures richer in furfural  $\log \gamma_1$  becomes negative. Mains stated that his results in this region were inexact because of slight fractionation of the vapours.

Removal of this fractionation would diminish the value of  $\frac{y_1}{x_1}$  in the equation for the calculation of activity, so that the maximum value of  $\log \gamma_1$  would appear at slightly lower molar fractions of furfural. The maximum would not be eliminated. Activity coefficients calculated from Mains' work are given in columns 1 to 3 of Table 3 and values of  $\log \gamma_1$  and  $\log \gamma_2$  are plotted in Figure 9.

This type of curve may then come partly from the assumption made in calculating the activities, that the simple gas laws apply to these vapours under normal and increased pressures. This assumption is hardly warranted, although no data are available concerning the behaviour of furfural vapour to give the extent of departure from the simple laws. It is much more probable, however, that the appearance of these maxima are related to the formation of strong hydrogen bridges between furfural and water molecules. Whatever the explanation it seems reasonable to conclude that activity coefficients obtained in this way may be graphed, and corrected activity coefficients obtained by smoothing the curve, but they cannot be used to correct the general trend of the  $\log \gamma$ -composition curve. It would be interesting to discover to what extent maxima observed in other systems, for example, chloroform and alcohol, depend on these two factors.

It is concluded that the data presented are reliable for mixtures with a mole fraction of furfural less than 0.6 (89 per cent. furfural). This is the region of interest for the production of furfural. Above 0.6 the experimental technique is difficult for the reasons given and the results are less accurate, though giving an essentially true picture of the equilibrium diagram.

## PART II.—THE INFLUENCE OF ADDED SALTS UPON THE MUTUAL SOLUBILITIES OF FURFURAL AND WATER

Sulphuric acid being the catalyst used in the commercial production of furfural, our investigation of added salts was restricted to metallic sulphates and those of sodium, magnesium, zinc, and aluminium were examined, and also potassium alum.

The method used was the determination of the miscibility temperature of furfural with varying amounts of aqueous solutions of the sulphates. The mixtures were sealed in pyrex bulbs of about 1.2 cc. capacity and heated and shaken in a large bath of dibutyl phthalate. The temperature of complete miscibility was taken as when the liquid appeared completely homogeneous and when all turbidity had disappeared.

The sulphates used were of A.R. quality. Furfural was purified according to the method of Evans and Aylesworth. The thermometers used were graduated in tenths of degrees and had been recently calibrated.

First, the solubility curve of furfural and water was redetermined and here we closely confirmed the findings of Evans and Aylesworth and not the earlier

work of Rothmund(14). Our results with the sulphates are reported in Tables 4 and 5, and are shown graphically for sodium and magnesium sulphates in Figures 10 and 11. Data for potash alum are not recorded because hydrolysis, with precipitation of a solid, took place rapidly. In no case was it possible to complete the solubility curves, for always at higher furfural proportions, crystallization of the salt took place before miscibility was attained. The miscibility temperature

TABLE 4  
FURFURAL AND SODIUM SULPHATE SOLUTIONS

With 1.168 w/v % Na <sub>2</sub> SO <sub>4</sub> Solution		With 5.016 w/v % Na <sub>2</sub> SO <sub>4</sub> Solution		With 10.00 w/v % Na <sub>2</sub> SO <sub>4</sub> Solution	
Furfural (%)	Miscibility Temperature (° C.)	Furfural (%)	Miscibility Temperature (° C.)	Furfural (%)	Miscibility Temperature (° C.)
10.21	63.9	7.78	64.6	4.56	52.0
11.13	71.0	9.87	90.2	7.27	101.8
12.74	82.1	11.00	100.2	11.72	134.7
18.76	106.2	13.56	112.2	13.82	146.0
21.56	111.9	16.48	124.8	15.72	156.0
22.65	115.5	20.19	135.1	16.94	159.2
25.55	118.9	24.24	142.7	19.70	169.3
28.15	122.4	26.54	146.6	20.85	173.0
33.72	127.1	29.24	151.2	23.25	176.6
35.95	128.4	32.00	154.6	25.65	184.0
38.06	129.3	35.41	158.6	28.00	190.5†
44.8	131.8	37.30	160.8	29.50	194.0†
48.3	133.2	38.70	163.1	30.0	195.5†
51.6	134.6	41.06	167.4		
56.9	138.5	42.30	171.2		
58.9	140.8	42.76	173.3		
62.9	148.7	44.60	185.7		
64.0	153.2*	45.60	—**		
67.2	159.6*				
71.0	169.9*				

\* Sodium sulphate crystallized and the results are untrustworthy.

\*\* Miscibility had not been attained at 222° C. when crystallization occurred.

† Decomposition with colouring made observation difficult.

is increased by the addition of a metallic sulphate. This increase, at first large, falls and then with increasing proportions of furfural in the mixture rises again rapidly, so that at some point the miscibility curve shows an inflection. This was observed with all sulphates, but with magnesium sulphate the inflection becomes a discontinuity. Thus, with 4.38 per cent. aqueous magnesium sulphate the miscibility temperature with 31.45 per cent. furfural is 145.8° C.; with 31.46 per cent. furfural it is 169° C. We have observed this very abrupt change

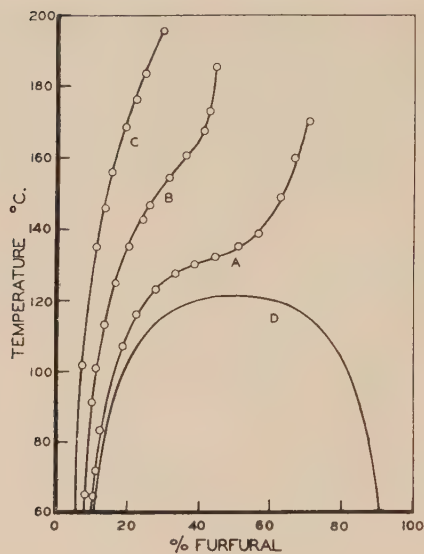


Fig. 10

- A. Furfural + 1.168 per cent.  $\text{Na}_2\text{SO}_4$  solution.  
 B. Furfural + 5.016 per cent.  $\text{Na}_2\text{SO}_4$  solution.  
 C. Furfural + 10.00 per cent.  $\text{Na}_2\text{SO}_4$  solution.  
 D. Furfural + distilled water.

repeatedly and suspect that it is due to a change of phase in the magnesium sulphate-water system. If this is so, then the inflections observed in the miscibility curves with other sulphates probably have a similar explanation.

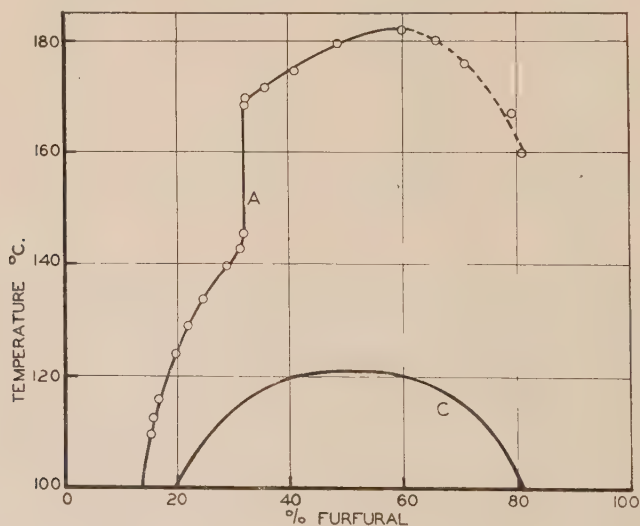


Fig. 11

- A. Furfural + 4.383 per cent.  $\text{MgSO}_4$  solution.  
 C. Furfural + distilled water.



The influence of added salts on the liquid-vapour diagram of furfural-water can be determined by superimposing the miscibility curves on the equilibrium curves of Figures 3, 4, and 5. The miscibility curve for 10 per cent. sodium sulphate solution cuts the liquidus curve at the points marked A in those figures (at 16 per cent. furfural for 9.83 kg. per cm.<sup>2</sup> pressure, 19 per cent. at 7.85 kg., and 22 per cent. at 6.17 kg.). Addition of 10 per cent. of sodium sulphate would therefore move the liquidus curves down to the azeotropic temperature at

TABLE 5  
FURFURAL AND MAGNESIUM SULPHATE SOLUTIONS

With 4.383% w/v MgSO <sub>4</sub> Solution			
Furfural (%)	Miscibility Temperature (° C.)	Furfural (%)	Miscibility Temperature (° C.)
8.35	73.8	31.68	169-170
11.52	97.7	34.4	170.8*
14.54	110.0	35.2	171.6*
16.24	116.2	37.8	172*
18.81	124.1	41.3	175
21.50	129.1	48.35	179.5-180
24.26	134.3	59.60	[182]
28.00	139.7	65.6	[180]
29.95	142.3	70.70	[176]
30.70	143.3	With 5.02% MgSO <sub>4</sub> soln.	
31.26	144.2	17.11	123.1
31.45	145.8	25.25	139.9
31.46	169.0	33.62	172
31.52	168.5		

\* For 32 per cent. and higher furfural contents the miscibility temperature was complicated by the fact that after clarification at about 170° C. a second liquid phase reappeared at 174°-176° C. The bracketed figures assigned to mixtures of more than 60 per cent. furfural content indicate that miscibility was not attained but magnesium sulphate crystals separated.

approximately these compositions. The change in the liquidus curve is seen to be very small, but increases a little with decreasing pressure on the system. The miscibility curve for 5 per cent. sodium sulphate solution cuts the liquidus curve on the furfural side of the azeotropic composition at all pressures above 6.17 kg.

In the commercial production of furfural under pressures of 7.0 to 10.5 kg. per cm.<sup>2</sup>, the condensate contains 4 to 6 per cent. of furfural which, from Figures 4 and 5, is seen to correspond to 1 or 2 per cent. of furfural in the liquid phase.

So it can be concluded that the addition either of sodium sulphate, or probably of any metallic sulphate, should be without influence on the elimination of furfural, except as it influences the formation of furfural from pentosans.

Simple sugars have but little effect on the mutual solubility curve. When 50.8 per cent. of furfural was added to 49.2 per cent. of a 5.1 per cent. aqueous solution of xylose, the miscibility temperature was 122.7° C. or 1.8° C. higher than the critical solution temperature of water and furfural.

TABLE 6  
FURFURAL AND ALUMINIUM AND ZINC SULPHATE SOLUTIONS

With 4.836% w/v Al <sub>2</sub> (SO <sub>4</sub> ) <sub>3</sub> Solution		With 5.00% w/v ZnSO <sub>4</sub> Solution	
Furfural (%)	Miscibility Temperature (° C.)	Furfural (%)	Miscibility Temperature (° C.)
10.52	83.3	14.37	104.2
12.42	92.3	22.00	138.3
14.31	101.4	38.50	164-165†
17.61	112.7	45.65	168-169†
19.30	116.1	71.15	167†
24.20	125.3		
31.55	134.0		
34.55	136.9		
37.90	139.4		
41.40	143.3		
44.00	159-160*		

\* At this temperature miscibility was almost achieved when crystallization ensued.

† At these temperatures miscibility appeared to be attained but shortly afterwards crystallization ensued.

### PART III.—THE DECOMPOSITION OF FURFURAL AT 180° C. IN AQUEOUS AND IN DILUTE SULPHURIC ACID SOLUTIONS

#### (a) *Experimental*

The apparatus used was essentially the same as that described in Part II. Mixtures (0.5 to 0.6 g.) of furfural and water or of furfural and dilute sulphuric acid were sealed in small pyrex bulbs, heated for the desired time at 180° C., rapidly cooled in air and in water and then placed in a small thick-walled brass cylinder with a funnel-shaped opening below. The bulb was crushed by a plunger and the contents were washed with a concentrated brine into a flask, the liquors carefully neutralized, and then furfural driven off in 200 ml. of aqueous distillate. It was estimated in the distillate by the thiobarbituric acid method(6).

It was necessary to steam distil the furfural to rid it of soluble products of higher molecular weight which might react with the thiobarbituric acid.

For reactions of furfural with water alone, doubly distilled water was freshly distilled in pyrex from potassium permanganate, and freshly distilled pure furfural was used. Even so, very erratic results were obtained until the reaction tubes were filled and sealed in an atmosphere of nitrogen.

TABLE 7  
REACTION BETWEEN FURFURAL AND WATER AT 180° C.

Furfural Concn. (wt. %)	Time (min.)	Unchanged Furfural (%)	$K \times 10^3$ 1st Order	$K \times 10^5$ 2nd Order	$K \times 10^{23}$ 11th Order
12.92	10	97.46	11.1	26.0	2.9
13.52	20	96.70	3.5	8.1	1.08
14.28	30	95.98	3.2	7.8	1.07
15.60	60	94.17	2.8	6.7	1.06
14.20	120	91.97	1.7	4.2	0.81
13.73	180	89.87	1.7	4.2	1.00

TABLE 8  
REACTION BETWEEN FURFURAL (4.52% w/w) AND 0.943% (0.192 N)  
AQUEOUS SULPHURIC ACID

Time (min.)	Unchanged Furfural (%)	$K = \frac{1}{t} \ln \frac{a-x_1}{a-x_2}$
10	90.5	0.00434
22	82.9	0.00317
30	77.8	0.00345
40	72.6	0.00301
60	61.8	0.00350
90	50.3	0.00298

Mean value of  $K = 0.00322 = 0.00742 \text{ min.}^{-1}$

### (b) Results and Discussion

Results of typical experiments are shown in Tables 7, 8, and 9. Pure furfural is seen to decompose only very slowly in pure water. There is appreciable reaction in the first ten minutes, presumably caused by traces of oxygen still present. Once this is over, the reaction is slow and apparently of very high order.

In the presence of dilute sulphuric acid furfural disappears rapidly and, as shown in Tables 8 and 9, the reaction is of the first order with regard to furfural. This is further emphasized by the close agreement in the values of  $K$  in Tables 9A and 9B where different concentrations of furfural are used with the same concentration of sulphuric acid.

TABLE 9A  
FURFURAL (3.81% w/w) AND 3.27% (0.667 N) AQUEOUS SULPHURIC  
ACID AT 180° C.

Time (min.)	Unchanged Furfural (%)	$K = \frac{1}{t} \ln \frac{a-x_1}{a-x_2}$
5	86.5	0.01260
10	74.0	0.01356
20	55.5	0.01249
30	38.3	0.01611
40	30.7	0.00961
60	16.7	0.01322
80	7.9	0.01625

Mean value of  $K = 0.01341 = 0.0309 \text{ min.}^{-1}$

TABLE 9B  
FURFURAL (13.4% w/w) AND 3.27% (0.667 N) AQUEOUS SULPHURIC  
ACID AT 180° C.

Time (min.)	Unchanged Furfural (%)	$K = \frac{1}{t} \ln \frac{a-x_1}{a-x_2}$
5	87.4	0.01170
10	76.5	0.01156
15	64.8	0.01442
20	56.0	0.01268
30	40.1	0.01451
40	30.3	0.01217
60	17.5	0.01192
75	9.8	0.01679

Mean value of  $K = 0.01322 = 0.0305 \text{ min.}^{-1}$

Over the range of acid concentrations studied (which covers those used in commercial production), the rate of disappearance of furfural is also roughly proportional to the acid concentration. Using the data of Table 8 to calculate the value of  $K$  in Table 9 a value of 0.0113 is obtained for an observed 0.0133.

These data can be used to calculate approximately the percentage of furfural lost by such reactions in commercial practice. Experiment B66 of Brown, Symons, and Wilson(15) can be used as an example. For reaction with 1.5 per cent. sulphuric acid the velocity constant is calculated as 0.0130. From Figure 5 of these authors the mean concentration of furfural in the distillate is 4 per cent. for the reaction period of 3 hours. Since it is practically certain that no fractional distillation occurred, the concentration of furfural in the liquid phase can be found from our liquid-vapour equilibrium curve (Fig. 5). This is seen to be 1 per cent. With a liquid-solid ratio of 1.5 the amount of liquid furfural in the digester is then 1.5 lb. per 100 lb. of oat-hulls. It follows that in 3 hours of reaction the amount of furfural lost is given by  $dx=0.0130(a-x)dt$  (where  $a-x$  is constant at 1.5 lb. and  $dt=180$  minutes) and amounts to 3.5 lb. In the experiment considered, the theoretical yield was 19 lb. and the loss 9.1 lb. So that some 18 per cent. of the furfural has been lost in this way, or 38 per cent. of all losses. While, therefore, much of the loss must come from destruction of pentosan and pentose by side reactions, or by reaction of furfural with other chemical compounds, this loss by decomposition of furfural is still considerable and the need to find a method for more rapid removal of furfural from the reaction mixture important. From the aspect of this one loss alone, a decrease of liquid-solid ratio appears desirable.

#### ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.

We have pleasure in thanking Mr. I. Brown of this Division for help and advice given in the course of this work, and particularly during the construction of the equilibrium still.

#### REFERENCES

- (1) MAINS, G. H.—*Chem. Metall. Engng.* **26** : 779 (1922).
- (2) EVANS, W. V., and AYLESWORTH, M. B.—*Ind. Engng. Chem.* **18** : 24 (1926).
- (3) DUNLOP, A. P., and PETERS, F. N.—*Ibid.* **32** : 1639 (1940).
- (4) LAZAREV, A. J.—*Org. Chem. Industr. U.S.S.R.* **6** : 258 (1939).
- (5) GRISWOLD, J., ANDRES, D., and KLEIN, V. A.—*Trans. Amer. Inst. Chem. Engrs.* **39** : 223 (1943).
- (6) MACKNEY, A. W., and REYNOLDS, T. M.—*J. Coun. Sci. Industr. Res. Aust.* **11** : 333 (1938).
- (7) DESNER, G. Y.—*J. Chem. Industr. (Moscow)* **7** : 695 (1930). (*Chem. Abstr.* **21** : 5817 (1931).)
- (8) LANDRIEU, P., BAYLOCQ, F., and JOHNSON, J. R.—*Bull. Soc. Chim. Fr.* (4)**45** : 36 (1929).
- (9) WALDEN, P.—*Z. Phys. Chem.* **70** : 574 (1910).
- (10) TIMMERMANNS, J.—*Bull. Soc. Chim. Belg.* **31** : 389 (1922).
- (11) STULL, D. R.—*Ind. Engng. Chem.* **39** : 517 (1947).
- (12) PRIGOGINE, I.—*J. Phys. Radium* (8)**5** : 185 (1944).
- (13) CARLSON, H. C., and COLBURN, A. P.—*Ind. Engng. Chem.* **34** : 581 (1942).
- (14) ROTHMUND, V.—*Z. Phys. Chem.* **26** : 454 (1898).
- (15) BROWN, I., SYMONS, E. F., and WILSON, B. W.—*J. Coun. Sci. Industr. Res. Aust.* **20** : 225 (1947).



# ALKYLENE TRITHIOCARBONATES AND ATTEMPTED SYNTHESSES OF 2,3-DIMERCAPTOPROPANOL (B.A.L.)

By C. C. J. CULVENOR\* and W. DAVIES\*

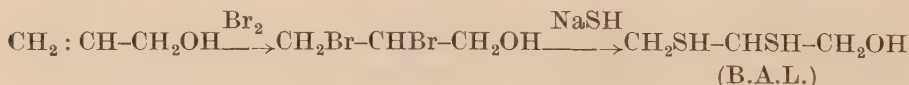
[Manuscript received February 9, 1948]

## Summary

Alkylene trithiocarbonates are usually conveniently made from the corresponding alkylene oxide or sulphide by the action of potassium xanthate or sodium trithiocarbonate. *Propylene* and *isobutylene trithiocarbonates* have been prepared, and it is found that the alkaline hydrolysis of alkylene trithiocarbonates to give the corresponding 1,2-dithiol is a general one. This hydrolysis should produce 2,3-dimercaptopropanol when applied to the trithiocarbonates derived from epichlorhydrin, glycidol and its acetate. However, these reactive substances do not produce pure trithiocarbonates. *Hydroxypropylene sulphide* has been obtained by treating chlorpropylene sulphide with sodium carbonate.

## I. INTRODUCTION

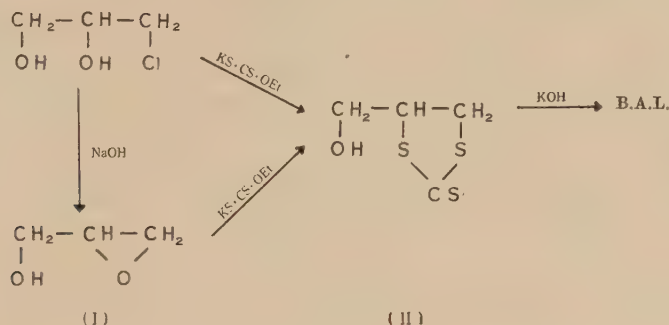
During the recent war, one of us examined methods of preparing 2,3-dimercaptopropanol, or British Anti-Lewisite (B.A.L.). This substance is invaluable for combating poisoning by arsenic. The general method of preparation(1, 2) was by the addition of halogen to allyl alcohol, followed by the action of alkali hydrosulphides, e.g.



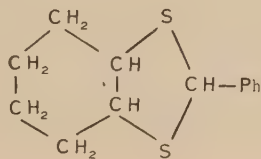
It was noticed that every specimen of B.A.L., whether made in England or Australia, contained halogen. The halogen-containing impurity could be almost completely removed by careful fractionation under reduced pressure, but this process, however, largely decomposed the B.A.L. Moreover, it was reported that specimens of B.A.L. could be dangerously toxic, and that the toxicity varied with different specimens. Now, likely impurities in the above method of preparation would be  $\text{CH}_2\text{SH}-\text{CHBr}-\text{CH}_2\text{OH}$  or  $\text{CH}_2\text{Br}-\text{CHSH}-\text{CH}_2\text{OH}$ , in which the halogen atom is beta to the S-atom. Some of the most toxic chemicals, e.g. "mustard gas", have this relationship, and, moreover, Bennett(3) has shown that the closely related 2-chlorethanethiol,  $\text{CH}_2\text{Cl}-\text{CH}_2\text{SH}$ , is also toxic. Accordingly, it seemed probable that the toxicity of B.A.L. would be reduced if it could be made in such a way that no halogen-containing impurities could be present.

\* Organic Chemistry Laboratory, University of Melbourne.

The hydrolysis of 3-hydroxypropylene trithiocarbonate (II) offered such a method, and the following synthesis was attempted, but was unsuccessful:



(i) *The Preparation and Hydrolysis of Trithiocarbonates.*—The preparation of several crystalline trithiocarbonates by the action of potassium xanthate on the corresponding ethylene oxides, sulphides, or chlorhydrins, described by Culvenor, Davies, and Pausacker(4), has now been extended. *Propylene* and *isobutylene trithiocarbonates* are readily obtained from the oxides, and also *p*-tolyl-(2-hydroxy-3-chlorpropyl)-sulphone is converted to 3-(*p*-tolyl-sulphonyl)-*propylene trithiocarbonate*.



(III)

Frassetto(5) first showed that ethylene trithiocarbonates gave ethylene dimercaptan, and it has now been found that propylene and cyclohexene dimercaptans may also be prepared from the corresponding trithiocarbonates. The conversion of trithiocarbonates to dithiols by alcoholic alkali is thus a general reaction. Satisfactory yields are obtained from the trithiocarbonates of lower molecular weight, but cyclohexene trithiocarbonate, which is less soluble in alcohol, gives only a poor yield of *cyclohexene dimercaptan*, which is readily characterized by combination with benzaldehyde to form 2-phenyl-hexahydrobenz-1,3-dithiolane (III). Piperonal condenses similarly to give the corresponding 2-piperonyl-hexahydrobenz-1,3-dithiolane.

(ii) *Attempted Synthesis of B.A.L.*—The initial stage was the preparation of 3-hydroxypropylene trithiocarbonate (II). Unfortunately, however, the reaction of glycidol (I) with potassium xanthate is not straightforward and leads to polymeric products. It has not been possible to isolate hydroxypropylene trithiocarbonate, which is apparently a non-distillable oil. Glycerol monochlorhydrin gives a similar result, although it is less reactive towards xanthate.

Since the reaction of an ethylene sulphide with xanthate is simpler than that of the oxide in so far as no intermediate is formed, *hydroxypropylene sulphide* was tried. Dachlauer and Jackel(6) reported that this sulphide is formed from glycidol and thiourea, but Culvenor, Davies, and Pausacker(4) showed that the product is polymeric. Monomeric hydroxypropylene sulphide has now been prepared by shaking chlorpropylene sulphide with aqueous sodium carbonate. It also, however, gives only a polymeric product with xanthate.

It was hoped that the use of glycidol acetate would overcome the difficulty since acetylation would prevent the OH-group from taking part in other possible condensations, and the acetyl group would be easily removed in the second stage of the synthesis. For the preparation of glycidol acetate, the previously described method of treating epichlorhydrin with potassium acetate(7) was found very unsatisfactory, but it was obtained readily and in excellent yield by the ketene acetylation of glycidol. However, it is found that the acetyl group is removed by hydrolysis with extreme ease (within a few minutes by mere mixing with water), so that this oxide also does not yield a trithiocarbonate.

Attempts were then made to prepare chlorpropylene trithiocarbonate with the expectation that, since the chlorine in chlorpropylene sulphide is hydrolysed at room temperatures by aqueous sodium carbonate, the stronger KOH would produce B.A.L. free from chlorine. Once more, however, only intractable yellow oils resulted from the interaction of xanthate with epichlorhydrin, glycerol dichlorhydrin, or chlorpropylene sulphide. With a reaction temperature of below 0° C., epichlorhydrin forms a fairly mobile product, but this cannot be purified.

The crude, partly polymerized products from glycidol and xanthate were submitted to treatment with alkali, but when no B.A.L. could be isolated from the reaction this method of synthesis was abandoned.

## II. EXPERIMENTAL

(i) *Propylene Trithiocarbonate.*—Propylene oxide (5 g.) is added to a xanthate solution prepared from KOH (10 g.) and CS<sub>2</sub> (18 g.) in alcohol (35 ml.), and cooling applied to keep the temperature below about 40° C. After standing for several hours, the yellow oil which separates is extracted with chloroform. It froths badly on heating under vacuum but distils at 136° C./0.2 mm. Yield, 9.0 g.; equivalent to 70% theory. (Found: S, 62.9%. C<sub>4</sub>H<sub>6</sub>S<sub>3</sub> requires S, 64.0%.)

Propylene trithiocarbonate (10 g.) is allowed to stand with a solution of KOH (15 g.) in alcohol (70 ml.) in an incubator at 32° C. for 6 days. The alcohol is then removed over a water-bath, and the residue treated with water, acidified,

and steam distilled. The oil separating from the distillate is extracted and distilled to give propylene dimercaptan (1.8 g.; 26%), b.p. 151–153° C.

(ii) *Isobutylene Trithiocarbonate*.—No heat is evolved when isobutylene oxide or sulphide is added to a xanthate solution, but on standing for several days the yellow oily isobutylene trithiocarbonate settles out. It is separated as above and has b.p. 166°–168° C./25 mm. or 126° C./0.2 mm. (Found: S, 59.3%.  $C_5H_8S_3$  requires S, 58.5%.) It may also be obtained from the chlorhydrin as follows. Isobutylene chlorhydrin (14 g.), KOH (23.8 g.), and  $CS_2$  (34.3 g.) in MeOH (100 ml.) are refluxed for 8 hours, then the products worked up as before to give trithiocarbonate (4.8 g.; 21%), b.p. 162°–170° C./24 mm. The trithiocarbonate melts at about  $-6^\circ$  to  $-12^\circ$  C.

(iii) 3-(*p*-Tolyl-sulphonyl)-propylene Trithiocarbonate.—*p*-Tolyl-(2-hydroxy-3-chloropropyl)-sulphone (0.5 g.) is refluxed with a solution of KOH (0.35 g.) and  $CS_2$  (0.5 g.) in alcohol (10 ml.) for 40 minutes. On cooling and diluting with water, an oil separates which soon solidifies to a yellow solid, readily soluble in benzene and chloroform, not so readily in alcohol, and insoluble in petroleum ether. Recrystallization from alcohol gives a microcrystalline powder, m.p. 128° C. (Found: C, 42.9; H, 4.2%.  $C_{11}H_{12}O_2S_4$  requires C, 43.4; H, 4.0%.)

(iv) *Cyclohexene Dimercaptan*.—Cyclohexene trithiocarbonate (8.9 g.) and KOH (8.5 g.) in alcohol (60 ml.) are kept at 35° C. for 3 weeks, then the alcohol distilled off under vacuum and the residue acidified with dilute  $H_2SO_4$  and steam distilled. The dimercaptan separates from the distillate and is extracted with ether. Yield, 1 g.; b.p. 97° C./15 mm. (Found: C, 48.65; H, 7.6%.  $C_6H_{12}S_2$  requires C, 48.64; H, 8.1%.) Addition of a drop of conc. HCl to a drop each of cyclohexene dimercaptan and benzaldehyde gives a pasty mass of the derivative 2-phenyl-hexahydrobenz-1,3-dithiolane, which gives crystals from alcohol, m.p. 115.5° C. (Found: C, 65.7; H, 6.8%.  $C_{13}H_{16}S_2$  requires C, 66.1; H, 6.8%.) In the same way, but with warming, piperonal yields the 2-piperonyl derivative, m.p. 125° C. (Found: C, 60.3; H, 6.2%.  $C_{14}H_{16}O_2S_2$  requires C, 60.0; H, 5.8%.)

(v) *Hydroxypropylene Sulphide*.—Chlorpropylene sulphide (4.6 g.) and sodium carbonate (5.3 g.), mixed in water (50 ml.) and alcohol (25 ml.) and shaken at room temperature for 30 hours, give after a continuous extraction with ether and drying over sodium sulphate hydroxypropylene sulphide (1.5 g.), b.p. 90° C./20 mm. (Found: S, 35.7.  $C_3H_6OS$  requires S, 35.5%.) Only a trace of the monomeric sulphide is obtained if glycidol is treated with thiourea in the presence of sodium carbonate. On standing, even the pure sulphide polymerizes slowly to a very viscous oil.

(vi) *Glycidol Acetate*.—An excess of ketene is passed through pure glycidol in an absorption tube cooled to 0° C. in an ice-bath. After the requisite time, the liquid is distilled and the glycidol acetate collected at 83° C./38 mm. Yield, 90%. The product is immiscible with water, but dissolves on shaking for a few minutes because of hydrolysis. When shaken with slightly less than the theoretical amount of NaOH, it dissolves quickly to give a neutral solution.



## III. REFERENCES

- (1) STOCKEN, L. A.—*J. Chem. Soc.* **1947**: 592 (1947).
- (2) SJÖBERG, B.—*Ber. dtsh. chem. Ges.* **75B**: 13 (1942).
- (3) BENNETT, G. M.—*J. Chem. Soc.* **121**: 2145 (1922).
- (4) CULVENOR, C. C. J., DAVIES, W., and PAUSACKER, K. H.—*Ibid.* **1946**: 1050 (1946).
- (5) FRASSETTI, P.—*Ber. dtsh. chem. Ges.* **38**: 491 (1905).
- (6) DACHLAUER, K., and JACKEL, L.—D.R.P. 636,708 (Oct. 15, 1936); Friedländer XXIII, 35.
- (7) LEVENE, P. A., and WALTI, A.—*J. Biol. Chem.* **79**: 363 (1928).



## STUDIES ON THE LIGNIN OF *EUCALYPTUS REGNANS*

### II. THE NATURE OF THE HYDROXYL GROUPS AND THE PRESENCE OF THE CARBONYL GROUP IN THIOLIGNIN

By J. W. T. MEREWETHER\*

[*Manuscript received April 9, 1948*]

#### *Summary*

*E. regnans* thiolignin reacts with *p*-toluenesulphonyl chloride in pyridine to form a hexatosyl derivative; its trimethyl ether reacts likewise to form a tritosyl derivative. Both compounds still have a free hydroxyl group which can be acetylated. Similarly they yield a hexabenzoate and tribenzoate respectively by the Schotten-Baumann reaction, but in pyridine, thiolignin reacts with benzoyl chloride to give a heptabenzoate and trimethylthiolignin a tetrabenzoate. No reaction takes place when trimethyl thiolignin is treated with triphenylchloromethane in pyridine. The above data are interpreted as evidence that of the four alcoholic hydroxyl groups three are secondary and one tertiary.

With phenylhydrazine, thiolignin yields a phenylosazone; with *p*-nitrophenylhydrazine it yields a *p*-nitrophenylhydrazone. On the other hand, trimethylthiolignin does not react with phenylhydrazine, indicating the absence of non-enolizable carbonyl groups. Thiolignin condenses with benzaldehydes indicating the presence of an active methylene group. From this evidence it is deduced that the grouping  $-\text{CH}_2-\text{CO}-\text{CHOH}-$  is present.

#### I. INTRODUCTION

In a previous paper(1) the results of a preliminary examination of the thiolignin isolated from the sulphate pulping of *Eucalyptus regnans* were described. Analysis of this thiolignin and of several of its derivatives indicated an empirical formula of  $\text{C}_{68}\text{H}_{76}\text{O}_{25}\text{S}$  and showed the presence of one mercaptan group and six hydroxyls. Two of these could be methylated with diazomethane, at least one of them being enolic, and while the remaining four could readily be acetylated only three could be methylated. The object of the work described in the present paper was to determine whether one or both acidic hydroxyls were enolic, to obtain evidence for the presence of the carbonyl group, and to obtain further data on the reactions of the alcoholic hydroxyl groups.

In view of the presence of at least one enolic hydroxyl, thiolignin should react in its keto form with the usual reagents for ketones. This is an aspect of lignin chemistry which has received comparatively little attention. Most workers

\* Research Laboratories, Australian Paper Manufacturers Limited, Melbourne.

who have reported the presence of enols have confined themselves to the preparation of enolic derivatives such as the methyl ether or acetate. Furthermore, the presence of the carbonyl group in lignin preparations is a subject on which the evidence is by no means straightforward(2).

It has been suggested by Hägglund, Johnson, and Trygg(3) that native lignin contains the carbonyl group and that the formation of lignosulphonic acid in the sulphite pulping process consists of a reaction between the carbonyl group and the bisulphite of the pulping liquor. A similar mechanism was postulated by Ahlm(4) to explain the formation of thioglignin in the sulphate process, and by Brauns(5) in methanol lignin. Brauns found that when native lignin is treated with methyl alcohol-hydrochloric acid it takes up two methoxyl groups which can subsequently be removed by 72 per cent. sulphuric acid, and considered this to be an argument for the presence of a carbonyl group. In support of this he claimed to have prepared a phenylhydrazone from native lignin, but his evidence is rather slender, no nitrogen figures being given.

Similarly, condensation products with phenylhydrazine were reported by Ahlm and Brauns(6) for thioglycollic acid lignin and by Virasoro(7) for acetoacetic ester lignin, but in neither case were nitrogen figures given. Condensation products with both semicarbazide and phenylhydrazine were recorded by Friedrich and Diwald(8) for their methanol lignin but they were of rather an indefinite character, the analytical figures being difficult to interpret. Dorée and Barton-Wright(9) found evidence for the presence of two carbonyl groups in alkali lignin, and the presence of non-enolizable carbonyl groups in formic acid lignin has been claimed by Wright and Hibbert(10). The presence of non-enolizable carbonyl groups may also be inferred from the evidence of Powell and Whittaker(11) who found that after methylation with dimethyl sulphate, alkali lignin condensed with phenylhydrazine to yield a product containing 7.5 per cent. nitrogen. Freudenberg *et al.*(12) found that Freudenberg lignin and also spruce wood itself react with hydrazine, similar results being obtained by Hess and Heumann(13) for straw.

Turning now to the question of the nature of the alcoholic hydroxyls, the use of dimethyl sulphate and aqueous alkali to obtain complete methylation of hydroxyl groups has been widely applied in lignin chemistry, and there has been little record of its failure to achieve this purpose. For example, it has been successfully applied to the complete methylation of a spruce thioglignin(4) and spruce alkali lignins(14), compounds which would be expected to show some similarity in behaviour to *E. regnans* thioglignin. On the other hand, the evidence in the case of spruce Freudenberg ("Cuproxam") lignin is contradictory, for while Brauns and Hibbert(15) have reported complete methylation, Freudenberg, Sohns, and Janson(16) obtained results similar to those of Lahey and Merewether(1). Freudenberg considered that the hydroxyl group that could not be methylated was tertiary, and drew attention to the fact that the molecule contained a methyl group attached to carbon(17) inferring from this that the grouping  $\text{:C}(\text{CH}_3)\text{OH}$  was very probably present.

The presence of a C-methyl group was also reported for *E. regnans* thioglignin(1).

II. *E. REGNANS* THIOLIGNIN

To provide further evidence as to the nature of these hydroxyl groups *E. regnans* thiolignin was treated with *p*-toluenesulphonyl chloride in pyridine. The product still contained a free hydroxyl group, acetylation yielding an acetate (1.9 per cent. acetyl). Since the tosylated product contains a free hydroxyl group it is presumably hexatosylthiolignin, a fact which is supported by its methoxyl content of 12.6 per cent. (calculated 12.4 per cent.). On the other hand, the carbon and sulphur analyses gave 59.7 per cent. carbon and 9.1 per cent. sulphur instead of the calculated values of 58.8 and 10.0 per cent. respectively. Further purification yielded a product with the same figures. This discrepancy can possibly be explained when one takes into account the fact that one of the groups to be tosylated is a mercaptan. The reaction of sulphonyl chlorides with mercaptans is not generally a satisfactory method for the preparation of thiosulphonates(18), for, although yields of up to 25 per cent. have been obtained in the presence of pyridine(19), the chief products of the reaction are usually sulphinic acids and disulphides, presumably resulting from a secondary reaction between the thiosulphonate and unchanged mercaptan(20).

If this is so, then normal tosylation should take place without any side reactions if the thiolignin is first methylated with diazomethane, thus blocking the mercaptan group, and then treated with *p*-toluenesulphonyl chloride in pyridine. This was found to be the case, the product analysing well for tritosyltrimethylthiolignin, one hydroxyl again remaining free as shown by the formation of an acetate.

Benzoylation experiments were even more instructive. Complete acylation of both thiolignin and trimethylthiolignin took place in pyridine; in aqueous alkali one hydroxyl remained free, thiolignin yielding a hexabenzoate and trimethylthiolignin a tribenzoate.

From these data it may be inferred that of the four alcoholic hydroxyls one is tertiary. On treating trimethylthiolignin with triphenylchloromethane in pyridine no reaction took place, trimethylthiolignin being recovered unchanged. From this the absence of primary hydroxyls is inferred, and the other three must therefore be secondary.

Of the two acidic hydroxyls, a spot-test had shown the presence of a keto-enol. Non-enolizable carbonyl groups were shown to be absent by methylating thiolignin with diazomethane in order to block the enolic hydroxyls and treating the resultant trimethylthiolignin with phenylhydrazine. No reaction took place, trimethylthiolignin being recovered unchanged. When thiolignin was treated with phenylhydrazine a product was obtained with 3.6 per cent. nitrogen, corresponding to four nitrogen atoms. In view of the fact that there are two acidic hydroxyls it was at first thought that this product was a diphenylhydrazone, both hydroxyls being enolic. However, *p*-nitrophenylhydrazine yielded a product with only 3.0 per cent. nitrogen, corresponding to a mono-*p*-nitrophenylhydrazone, indicating the presence of only one carbonyl group, and that the reaction product with phenylhydrazine was actually a phenylosazone. From this it is deduced that the group  $\text{-CO-CHOH-}$  is present, and that the other acidic hydroxyl must be phenolic.



In view of the presence of an enolizable carbonyl group it was thought possible that there would be an active methylene group adjacent. Accordingly, thiolignin was condensed with benzaldehyde and *m*-nitrobenzaldehyde, yielding condensation products whose analysis showed that one molecule of aldehyde had reacted.

It is concluded, therefore, that the grouping  $-\text{CH}_2\text{CO-CHOH-}$  is present, and the formula for *E. regnans* thiolignin(1) may be expanded to

$\text{C}_{52}\text{H}_{35}\text{O}_{10}\cdot(\text{OCH}_3)_9\cdot\text{C}(\text{CH}_3)\text{OH}\cdot(\text{CHOH})_2\cdot(\text{CH}_2\text{CO}\cdot\text{CHOH})\cdot\text{OH}\cdot\text{SH}$ ,  
the last hydroxyl being phenolic.

### III. EXPERIMENTAL

#### (a) General Procedure

To avoid repetition, the following general procedure was used for the isolation and purification of derivatives. At the completion of a reaction the mixture was filtered and added dropwise with stirring to ten volumes of precipitating solvent, usually either water or ether. The crude product was filtered under suction, washed with fresh solvent, air-dried, and finally dried in a vacuum desiccator (15–25 mm.) over sulphuric acid. This was then dissolved in a suitable solvent to give a 10 per cent. w/v solution, filtered, and added dropwise with stirring to ten volumes of anhydrous ether. The precipitate was filtered under suction, washed with pure ether, air-dried, and finally dried in a vacuum desiccator over sulphuric acid. Prior to analysis a sample was dried over phosphorus pentoxide at 57° C. and 15–25 mm. in an Abderhalden drying pistol. Purification was repeated to a constant methoxyl content.

#### (b) Hexatosylthiolignin

Thiolignin (5 g.) (1) in anhydrous pyridine (50 cc.) was reacted with redistilled *p*-toluenesulphonyl chloride (12.5 g.) for 24 hours at room temperature. The crude product, precipitated by pouring into water, weighed 8.65 g. Yield after two purifications from chloroform-ether, 7.0 g.

Found: C, 59.7; H, 5.1; S, 9.1;  $\text{OCH}_3$ , 12.6%.

Calculated for  $\text{C}_{110}\text{H}_{112}\text{O}_{37}\text{S}_7$  (2248): C, 58.8; H, 5.0; S, 10.0;  $\text{OCH}_3$ , 12.4%.

Hexatosylthiolignin is a brown powder which begins to soften at 200°–205° C., is soluble in chloroform and pyridine, slightly soluble in acetone, and insoluble in water, dilute acids and alkalis, ether, alcohol, benzene, petroleum ether, acetic acid, ethyl acetate, and dioxan.

#### (c) Acetylhexatosylthiolignin

Hexatosylthiolignin (7.9 g.) in anhydrous pyridine (40 cc.) was reacted with acetic anhydride (28 cc.) for 48 hours at room temperature. The crude product isolated from ether weighed 7.6 g. Yield after two purifications from chloroform-ether 6.3 g.

Found: C, 58.8; H, 5.1; S, 9.4;  $\text{OCH}_3$ , 11.7;  $\text{COCH}_3$ , 2.3%.

Calculated for  $C_{112}H_{114}O_{38}S_7$  (2290): C, 58.8; H, 5.0; S, 9.8;  $OCH_3$ , 12.2;  $COCH_3$ , 1.9%.

Acetylhexatosylthiolignin is a brown powder which commences to soften at 182°–186° C. and is soluble in the same solvents as hexatosylthiolignin.

(d) *Tritosyltrimethylthiolignin*

Trimethylthiolignin (5 g.), prepared as previously described(1), was tosylated under the conditions used for the tosylation of thiolignin. Yield 6.05 g. Yield after purification from chloroform-ether 5.3 g.

Found: C, 60.6; H, 5.6; S, 6.8;  $OCH_3$ , 20.0%.

Calculated for  $C_{92}H_{106}O_{31}S_4$  (1830): C, 60.4; H, 5.5; S, 7.0;  $OCH_3$ , 20.3%.

Tritosyltrimethylthiolignin is a brown powder which commences to soften at 158°–160° C., is soluble in acetone, chloroform, glacial acetic acid, and pyridine, slightly soluble in ethyl acetate, and insoluble in water, dilute acids and alkalis, alcohol, ether, benzene, petroleum ether, and dioxan.

(e) *Acetyltritosyltrimethylthiolignin*

Tritosyltrimethylthiolignin (5.7 g.) was acetylated under the conditions used for the acetylation of hexatosylthiolignin. Yield 5.5 g. Yield after two purifications from chloroform-ether 4.3 g.

Found: C, 60.0; H, 5.6; S, 6.5;  $OCH_3$ , 19.7;  $COCH_3$ , 1.9%.

Calculated for  $C_{94}H_{102}O_{32}S_4$  (1872): C, 60.3; H, 5.5; S, 6.9;  $OCH_3$ , 19.9;  $COCH_3$ , 2.3%.

Acetyltritosyltrimethylthiolignin is a brown powder which commences to soften at 168°–170° C. and has the same solubility characteristics as tritosyltrimethylthiolignin.

(f) *Heptabenzoylthiolignin*

Thiolignin (2 g.) was dissolved in anhydrous pyridine (20 cc.) and freshly redistilled benzoyl chloride (5.5 cc.) added. The mixture was allowed to stand for 48 hours, after which it was poured on to crushed ice. The product separated out as a tarry mass, so the mixture was extracted with chloroform, the chloroform solution washed successively with water, aqueous sodium bicarbonate, and again with water, and then dried overnight over anhydrous magnesium sulphate. It was then concentrated down to about 25–30 cc. and the product isolated from ether. Yield 1.5 g. Yield after two purifications from chloroform-ether 1.3 g.

Found: C, 68.5; H, 5.0; S, 1.7;  $OCH_3$ , 13.5;  $COC_6H_5$ , 35.9%.

Calculated for  $C_{117}H_{104}O_{32}S$  (2052): C, 68.4; H, 5.1; S, 1.6;  $OCH_3$ , 13.5;  $COC_6H_5$ , 35.9%.

Heptabenzoylthiolignin is a brown powder which commences to darken in colour at 190°–192° C. and to soften at 208°–210° C. It is soluble in acetone, chloroform, benzene, pyridine, and dioxan, and insoluble in water, dilute acids and alkalis, alcohol, ether, petroleum ether, glacial acetic acid, and ethyl acetate.

(g) *Tetrabenzoyltrimethylthiolignin*

Trimethylthiolignin (2 g.) was benzoylated under the conditions described in the previous experiment. Yield 1.6 g. Yield after two purifications from chloroform-ether 1.5 g.



Found : C, 66.7 ; H, 5.5 ; S, 2.0 ;  $\text{OCH}_3$ , 20.8 ;  $\text{COC}_6\text{H}_5$ , 23.7%.

Calculated for  $\text{C}_{99}\text{H}_{98}\text{O}_{29}\text{S}$  (1782) : C, 66.6 ; H, 5.5 ; S, 1.8 ;  $\text{OCH}_3$ , 20.9 ;  $\text{COC}_6\text{H}_5$ , 23.6%.

Tetrabenzoyltrimethylthiolignin is a buff powder which commences to soften at  $183^\circ\text{--}185^\circ\text{C}$ ., is soluble in acetone, chloroform, glacial acetic acid, ethyl acetate, benzene, pyridine, and dioxan, and insoluble in water, dilute acids and alkalis, alcohol, ether, and petroleum ether.

(h) *Hexabenzoylthiolignin*

Thiolignin (3 g.) was dissolved in 10 per cent. aqueous sodium hydroxide (30 cc.) and redistilled benzoyl chloride (3 cc.) added. The mixture was then shaken vigorously. After a few minutes no smell of benzoyl chloride could be detected and the mixture was quite warm. Further quantities of benzoyl chloride were added, 1 cc. at a time, shaking well between additions, and adding more 10 per cent. NaOH when necessary. After an additional 7 cc. benzoyl chloride had been added the odour of benzoyl chloride persisted and the product had separated out as a solid cake. The mixture was allowed to stand for an hour on the side of a water-bath, after which the product was filtered under suction and washed well with water. It was purified once from dioxan-water and once from dioxan-ether. Yield 3.8 g.

Found : C, 67.6 ; H, 5.3 ; S, 1.5 ;  $\text{OCH}_3$ , 14.0 ;  $\text{COC}_6\text{H}_5$ , 31.8%.

Calculated for  $\text{C}_{110}\text{H}_{100}\text{O}_{31}\text{S}$  (1948) : C, 67.8 ; H, 5.1 ; S, 1.6 ;  $\text{OCH}_3$ , 14.3 ;  $\text{COC}_6\text{H}_5$ , 32.4%.

Hexabenzoylthiolignin is a brown powder which commences to soften at  $188^\circ\text{--}190^\circ\text{C}$ ., is soluble in acetone, chloroform, benzene, pyridine, dioxan, and ethyl acetate, is slightly soluble in glacial acetic acid, and insoluble in water, dilute acids and alkalis, alcohol, ether, and petroleum ether.

(i) *Tribenzoyltrimethylthiolignin*

Trimethylthiolignin (3 g.) was benzoylated as described in the preceding experiment. Yield 3.5 g.

Found : C, 66.0 ; H, 5.4 ; S, 1.9 ;  $\text{OCH}_3$ , 22.0 ;  $\text{COC}_6\text{H}_5$ , 19.0%.

Calculated for  $\text{C}_{92}\text{H}_{94}\text{O}_{28}\text{S}$  (1678) : C, 65.9 ; H, 5.6 ; S, 1.9 ;  $\text{OCH}_3$ , 22.2 ;  $\text{COC}_6\text{H}_5$ , 18.9%.

Tribenzoyltrimethylthiolignin is a brown powder which commences to soften at  $189^\circ\text{--}193^\circ\text{C}$ ., is soluble in acetone, chloroform, glacial acetic acid, ethyl acetate, benzene, pyridine, and dioxan, and insoluble in water, dilute acids and alkalis, alcohol, ether, and petroleum ether.

(j) *Attempted Tritylation of Trimethylthiolignin*

Trimethylthiolignin (3 g.) was dissolved in dry pyridine (30 cc.) and triphenylchloromethane (5 g.) added. The mixture was kept on a water-bath for 48 hours, poured on to crushed ice, the product filtered, washed well with water, and air-dried. It was purified once from chloroform-ether. Yield 2.2 g.

Found :  $\text{OCH}_3$ , 26.6%.

Calculated for trimethylthiolignin :  $\text{OCH}_3$ , 27.2%.

(k) *Thiolignin Phenyllosazone*

Thiolignin (15 g.) in acetic acid (200 cc.) and dioxan (100 cc.) was reacted with crystalline phenylhydrazine (15 g.) for 24 hours on the water-bath. It was then concentrated down to 150 cc. and the crude product isolated from ether. Yield 16 g. Yield after two purifications from dioxan-ether 12.8 g.

Found: C, 64.3; H, 5.5; S, 2.1;  $\text{OCH}_3$ , 18.7; N, 3.6%.

Calculated for  $\text{C}_{80}\text{H}_{86}\text{N}_4\text{O}_{23}\text{S}$  (1502): C, 63.9; H, 5.7; S, 2.1;  $\text{OCH}_3$ , 18.6; N, 3.7%.

Thiolignin phenyllosazone is a dark brown powder that begins to soften at  $206^\circ\text{--}210^\circ\text{C.}$ , is soluble in dioxan, pyridine, and acetic acid, slightly soluble in acetone and ethyl acetate, and insoluble in water, alcohol, ether, chloroform, benzene, and petroleum ether.

(l) *Thiolignin p-Nitrophenylhydrazone*

Thiolignin (5 g.) in dioxan (50 cc.) was reacted with a solution of *p*-nitrophenylhydrazine (6 g.) in glacial acetic acid (60 cc.) for 4 hours on the water-bath. The mixture was then allowed to stand at room temperature for 24 hours and the product isolated and purified as described for the osazone. Final yield 4.2 g.

Found: C, 61.0; H, 5.4; S, 2.1;  $\text{OCH}_3$ , 19.2; N, 3.0%.

Calculated for  $\text{C}_{74}\text{H}_{81}\text{N}_3\text{O}_{26}\text{S}$  (1459): C, 60.9; H, 5.6; S, 2.2;  $\text{OCH}_3$ , 19.1; N, 2.9%.

Thiolignin *p*-nitrophenylhydrazone is a brown powder which begins to soften at  $210^\circ\text{--}212^\circ\text{C.}$ , is soluble in acetone, dioxan, and pyridine, slightly soluble in alcohol, chloroform, and acetic acid, and insoluble in water, ether, benzene, ethyl acetate, and petroleum ether.

(m) *Treatment of Trimethylthiolignin with Phenylhydrazine*

Trimethylthiolignin (1 g.) was treated with phenylhydrazine under the conditions used to prepare the osazone. The product was nitrogen free.

Found:  $\text{OCH}_3$ , 27.3%.

Calculated for trimethylthiolignin:  $\text{OCH}_3$ , 27.2%.

(n) *Benzalthiolignin*

Thiolignin (4 g.) was dissolved in 10 per cent. aqueous sodium hydroxide (40 cc.) and redistilled benzaldehyde (2 g.) added. The mixture was warmed on the water-bath overnight, after which it was cooled and acidified with 10 per cent. hydrochloric acid. The precipitate was filtered under suction, washed with water and partially air-dried. The moist product was dissolved in dioxan and the solution distilled until pure dioxan started to distil. On cooling, the solution was filtered from inorganic material, the volume adjusted to 40 cc. and the product isolated from ether (400 cc.). The crude product was purified again from dioxan-ether. Final yield 3.0 g.

Found: C, 63.4; H, 5.6; S, 2.4;  $\text{OCH}_3$ , 19.9%.

Calculated for  $\text{C}_{75}\text{H}_{80}\text{O}_{25}\text{S}$  (1412): C, 63.7; H, 5.7; S, 2.3;  $\text{OCH}_3$ , 19.8%.

Benzalthiolignin is a brown powder which begins to soften at 240°–245° C., is soluble in acetone, dioxan, and pyridine, slightly soluble in alcohol, glacial acetic acid, and ethyl acetate, and insoluble in water, chloroform, ether, petroleum ether, and benzene.

(o) *m*-Nitrobenzalthiolignin

Thiolignin (4 g.) was condensed with *m*-nitrobenzaldehyde (2.5 g.) under the conditions described for the preparation of benzalthiolignin. Final yield 3.0 g.

Found: C, 61.8; H, 5.3; S, 2.1; OCH<sub>3</sub>, 19.4; N, 1.1%.

Calculated for C<sub>75</sub>H<sub>79</sub>NO<sub>27</sub>S (1457): C, 61.8; H, 5.5; S, 2.2; OCH<sub>3</sub>, 19.2; N, 1.0%.

*m*-Nitrobenzalthiolignin is a brown powder which begins to soften at 213°–217° C. and is soluble in the same solvents as benzalthiolignin.

#### IV. ACKNOWLEDGMENTS

The author is indebted to Messrs. Australian Paper Manufacturers Limited for permission to publish this work; to Dr. F. N. Lahey, University of Melbourne, co-author of the first paper of this series, for his continued interest and advice; and to Miss Betty Hickox and Mr. N. Lottkowitz who carried out many of the analyses.

#### V. REFERENCES

- (1) LAHEY, F. N., and MEREWETHER, J. W. T.—*Aust. J. Sci. Res. A* **1**: 112 (1948).
- (2) PHILLIPS, M.—Wise's "Wood Chemistry", 1st Ed., p. 299. (Reinhold: New York, 1944.)
- (3) HÄGGLUND, E., JOHNSON, T., and TRYGG, L. H.—*Svensk PappTidn.* **32**: 815 (1929).
- (4) AHLM, C. E.—*Paper Tr. J.* **113**(13): 115 (1941).
- (5) BRAUNS, F. E.—*J. Amer. Chem. Soc.* **61**: 2120 (1939); *Paper Tr. J.* **108**(1): 42 (1939).
- (6) AHLM, C. E., and BRAUNS, F. E.—*J. Amer. Chem. Soc.* **61**: 277 (1939).
- (7) VIRASORO, E.—*An. Inst. Cient. Tecnol. Argentina* **12/13**: 175 (1943). (*Chem. Abstr.* **38**: 6553 (1944).)
- (8) FRIEDRICH, A., and DIWALD, J.—*Monatsh. Chem.* **46**: 31 (1925).
- (9) DORÉE, C., and BARTON-WRIGHT, E. C.—*Biochem. J.* **21**: 290 (1927).
- (10) WRIGHT, G. F., and HIBBERT, H.—*J. Amer. Chem. Soc.* **59**: 125 (1937).
- (11) POWELL, W. J., and WHITTAKER, H.—*J. Chem. Soc.* **125**: 357 (1924); *Ibid.* **127**: 132 (1925).
- (12) FREUDENBERG, K., ENGLER, F., FLICKINGER, E., SOBEK, A., and KLINK, F.—*Ber. dtsh. chem. Ges.* **71**: 1810 (1938).
- (13) HESS, K., and HEUMANN, K. E.—*Ibid.* **75**: 1802 (1942).
- (14) MARSHALL, H. B., BRAUNS, F. E., and HIBBERT, H.—*Canad. J. Res.* **13B**: 103 (1935).
- (15) BRAUNS, F. E., and HIBBERT, H.—*Ibid.* **13B**: 78 (1935).
- (16) FREUDENBERG, K., SOHNS, F., and JANSON, A.—*Liebigs Ann.* **518**: 62 (1935).
- (17) FREUDENBERG, K., and SOHNS, F.—*Ber. dtsh. chem. Ges.* **66**: 262 (1933).
- (18) CONNOR, R.—Gilman's "Organic Chemistry." 2nd Ed., p. 907. (John Wiley: New York, 1944.)
- (19) GIBSON, D. T., MILLER, C. J., and SMILES, S.—*J. Chem. Soc.* **127**: 1821 (1925).
- (20) PAULY, C., and OTTO, R.—*Ber. dtsh. chem. Ges.* **10**: 2181 (1877).

# THE SCATTERING OF FAST POSITRONS AND ELECTRONS IN GOLD (FOR THE RANGE 0.39–1.1 MEV.)

By W. B. LASICH\*

[*Manuscript received August 3, 1948*]

## *Summary*

The Wilson cloud chamber has been employed with the method of projected angles in a study of the scattering of positrons and electrons in thin gold foil, for the energy range 0.39–1.1 Mev. The small angle scattering is in reasonable agreement with current multiple scattering theory. Further, the investigation of the large angle scattering gives indications in favour of a spin orbital interaction, and the absolute intensity of the scattering for positrons and electrons under these conditions is believed to be approximately correct.

## I. INTRODUCTION

In considering the interaction of fast electrons with atoms, electron spin must be taken into account and Dirac's relativistic wave equations used in place of the ordinary Schroedinger equation. It then appears that the interaction of the spin and the orbital motion of the electron is most marked for encounters with atoms of the heavy elements. It is interesting, however, that for such atoms, the scattering predicted by theory is not the same for positrons as for electrons, and while recent experiments tend to confirm the theory for electrons, the effects are difficult to demonstrate with electrons alone. There has been no comparison of positron and electron scattering since the work of Fowler and Oppenheimer(1) in 1938.

Although these workers were successful in detecting a difference of the correct order, the observations were too few in number to have great significance. Further investigation should provide the most direct experimental evidence for the spin of the positron, as the existence of such spin cannot be inferred from spectroscopic data, or readily be demonstrated by the polarization of positron beams by double scattering.

In the present experiments the angular distribution of positrons and electrons scattered in a thin gold foil has been studied in a cloud chamber, the scattering for the energies involved being practically all elastic and due almost entirely to nuclear collisions. The angular distribution of particles is regarded as compounded of a multiple scattering distribution merging finally into a single scattering tail at very large angles.

\* C.S.I.R. Atomic Physics Section, Physics Department, University of Melbourne.



The treatment of multiple scattering from a foil was first given by Williams(2) assuming a Gaussian form, but a more complete analytical method was developed later by Goudsmit and Saunderson(3). Where single scattering conditions hold, the calculated intensity is more sensitive to any modification of Coulomb interaction, and distinctive curves may be drawn for positron and electron scattering.

The single scattering of electrons by a heavy element (Hg ;  $Z=80$ ) has been calculated by Bartlett and Watson(4), and extended by Massey(5) to positrons. The spin orbital interaction causes the intensity of scattering of electrons by a heavy element (e.g. gold) to be greater than the value given by the Rutherford

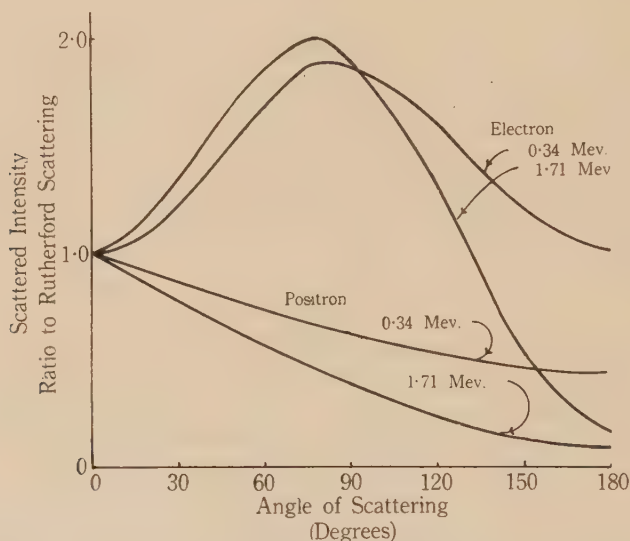


Fig. 1.—Calculated scattering *v.* angle for positrons and electrons.

formula, whilst that for positrons is smaller. The extent of the deviation increases with the energy of the particle. It also increases with the angle of scattering to about  $90^\circ$  in electrons, and continuously for all angles of scattering in positrons (Fig. 1).

## II. CLOUD CHAMBER

The magnetic field cloud chamber used was semi-automatic and 14 cm. in diameter, containing air at 60 cm. pressure and an alcohol-water vapour mixture. A single 35 mm. camera with an  $f/2.8$  lens was mounted with its axis vertical.

The source of positrons was an impure specimen of  $\text{Co}^{56}$  with a half-life of 72 days and an end-point of 1.2 Mev. It was mounted in a shallow cell having a thin celluloid window of area  $0.8 \text{ cm.}^2$ . Owing to the relative impurity of the radio-cobalt the maximum strength of the positron source was limited to about 60 microcuries of  $\gamma$ -ray activity. In view of this fact no attempt was made to select a narrow range of energies from outside the chamber, but a shaped



lead block was placed in the fringing field which imposed a lower limit of about 0.35 Mev. to the particles which entered. The lead block also served to shield the chamber from the  $\gamma$ -rays of the source.

For strictly comparable results with negatively charged particles, the same arrangement was used for the electron sources with reversal of the magnetic field. For the upper part of the energy range, the electron source was prepared from a drop of radio phosphorus in  $\text{H}_3\text{PO}_4$  on filter paper, and for the lower part  $\text{Ra}(\text{D}+\text{E})$  lead in the form of thin metal shavings. The gold foil was mounted on a light brass frame  $6 \times 2.2$  cm., with a layer either side of the frame to minimize turbulence effects in the gas. The outer surfaces were coated with about 2 mg. of smoked carbon black to eliminate scattering of light out of the plane layer. The particles were incident between  $60^\circ$  and  $110^\circ$  of the plane of the foil, making a total effective thickness ( $t$ ) of  $1.2 \times 10^{-4}$  cm. of gold.

Following the application of the magnetic field of 484 gauss the source was exposed by an electric shutter for the duration of the cloud chamber expansion. The particles were allowed to enter the chamber through a slot 1 cm. deep and 4.5 cm. long, ground in the glass wall of the chamber, and covered with a copper window of 0.004 cm. thickness mounted on a brass shoe having the same curvature as the chamber wall and fastened to the glass with picein.

The cloud chamber was illuminated in a plane layer 2 cm. deep, provided by six 500 w. horizontal filament tungsten lamps, beamed with cylindrical lenses. These were operated at a dull red heat, and flashed for about 0.15 second duration, with an input of 7 kilowatts. A photographic shutter was used to prevent "trailing" of the images of the cloud chamber tracks due to the "after-glow" of the lamps.

### III. THE PROJECTION METHOD

The projected angles of scattering in the horizontal plane were yielded by the camera on the vertical axis.

The experimental distributions were compared with various theoretical formulae transformed for plane projection and subjected to a criterion of minimum length of projected track following the treatment of Ó'Ceallaigh and MacCárthaigh(6).

The general law of scattering is of the form

$$P(\theta)d\theta = Ntg(Z,\beta) \cdot F(\theta) \cdot 2\pi \sin \theta \cdot d\theta \dots\dots\dots (1)$$

where  $P(\theta)d\theta$  is the probability of a particle moving with velocity  $\beta c$  through a medium containing  $N$  scattering nuclei, of atomic number  $Z$ , per unit volume, being scattered once through an angle lying between  $\theta$  and  $\theta + d\theta$ , in a path length of  $t$  cm. We take the boundaries of the illuminated layer to be the planes  $y = \pm a$ , and  $\lambda_c$  the minimum projected length of tracks to be included in the measurements. Taking  $\varphi$  as the projection of  $\theta$  in a plane perpendicular to the

axis of the camera ( $y$ -axis), the probability of scattering into the angular range  $\varphi$  to  $\varphi + d\varphi$  is

$$P(\varphi)d\varphi = 4Ntg(Z, \beta) \int_0^{\arccot \rho} F(\theta) [\cos \chi - \rho \sin \chi] d\chi \dots\dots\dots (2)$$

where  $\rho = \lambda_c/2a$ , and  $\cos \theta = \cos \varphi \cos \chi$ .

This formula is equivalent to that given by Ó'Ceallaigh and MacCárthaigh(6). Its derivation is discussed in the Appendix. It is assumed that the incident particles travel normal to the  $y$ -axis in a parallel beam of uniform intensity covering the entire depth of the illuminated layer. Under the experimental conditions the particles from the emitter suffer considerable scattering in the copper window, so that it becomes the virtual source and contributes to the uniformity of the particle density over the illuminated layer. The particles are incident on the gold foil making an angle, in general, of less than  $7^\circ$  with the horizontal plane.

The number of particles, which were actually incident on the foil yielding the observed distribution of scattered tracks of length greater than  $\lambda_c$ , is difficult to determine directly, since numerous tracks appeared to stop at the foil, making it somewhat difficult to decide in many cases whether the effect was due to a non-uniform expansion, or to the particle really being deflected steeply out of the illuminated layer. In practice then, a number of tracks appear to come close to the foil on the incident side, without trace of an emergent branch, and it was therefore decided to select only those scattering events with emergent branches of length  $\lambda_c$  and to make the theoretical comparison on this basis.

Let  $S$  be the ratio of the number of particles, which emerge with length  $\lambda_c$ , to the actual number of incident particles, then  $P'(\varphi) = \frac{1}{S} P(\varphi)$ , where  $P'(\varphi)$  is then the probability of  $\varphi$  being in the range  $\varphi$  to  $\varphi + d\varphi$ , for the events with  $\lambda \geq \lambda_c$ . Values of  $S$  are calculated in the Appendix.

#### IV. SELECTION AND MEASUREMENT OF TRACKS

The photographs were projected on to a measuring table with an overall magnification of 2.5 times natural dimensions, and the curvature of the projected tracks was measured by matching to a set of curves to cover the range 5 to 12 cm. in radius of curvature of cloud chamber track.

To reduce as far as possible selective errors in the measurement of tracks, a technique was used in which first of all the half of the photograph containing the incident tracks was examined, and only those were selected which were clearly visible to within a few millimetres of the foil. The emergent tracks of each photograph were examined in turn, and if they were likewise satisfactorily formed, with a length in each case exceeding 2 cm., the track was included in the measurements. In a few cases inelastic collisions were observed, but in general the curvature of the path on the two sides of the foil was the same within the accuracy of measurement, which was about  $\pm 1$  cm. in the radius of curvature. Similarly the estimate of the angle of scattering was reproducible to an accuracy of about  $\pm 2^\circ$ .

## V. COMPARISON OF RESULTS WITH THEORY

(i) *Scattering Formulae*.—The general form of the angular distribution of the scattered electrons has been calculated from the formulæ of Goudsmit and Saunderson(3), and in general Tallqvist's(7) tables of Legendre polynomials to the order  $n=32$  gave their function,  $f(\theta)$  to sufficient numerical accuracy for the calculation of  $P(\varphi)$ .

Where the exact nature of the interaction of the electron and nucleus is important in the single region of scattering, the theory of Goudsmit and Saunderson may be subject to some uncertainty, but the single scattering of

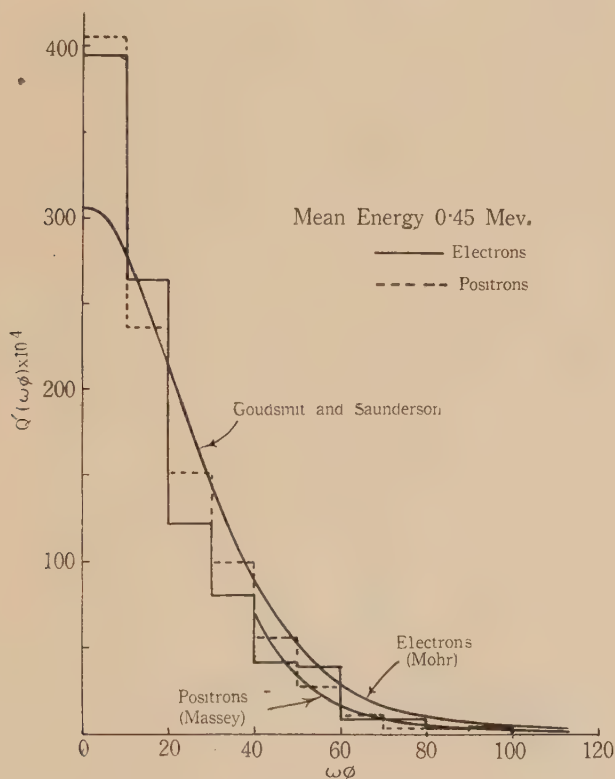


Fig. 2.—Angular distribution of scattering in the range 0.35–0.56 Mev.

electrons in gold has been specifically investigated by Mohr(8) who derives the formula

$$F(\theta) = \left( 1 + \frac{\pi}{2} \cdot \frac{Z}{137} \sin \theta/2 \right)^2 \operatorname{cosec}^4 \theta/2.$$

Similarly for positrons, Massey's results are closely expressed by the equation

$$F(\theta) = (1 + a \sin \theta/2 + b \sin^2 \theta/2) \operatorname{cosec}^4 \theta/2$$

where  $a$  and  $b$  are numerical coefficients chosen for each particular energy of the particle.

The method of projecting these formulae to obtain  $P(\varphi)$  has been published by Barker(9). The contribution of multiple scattering has been estimated with the procedure of Chase and Cox(10), who multiply the single scattering intensity by the factor  $1 + \epsilon^2(\text{cosec}^2 \theta/2 - \frac{1}{2})$ , where  $\epsilon^2$  is the mean square angle of deflection in the foil. However, the value of  $\epsilon^2$ , which was required to fuse the modified single curve of scattering to the Goudsmit and Saunderson curve at small angles, was found to be rather less than the value obtained by direct calculation.

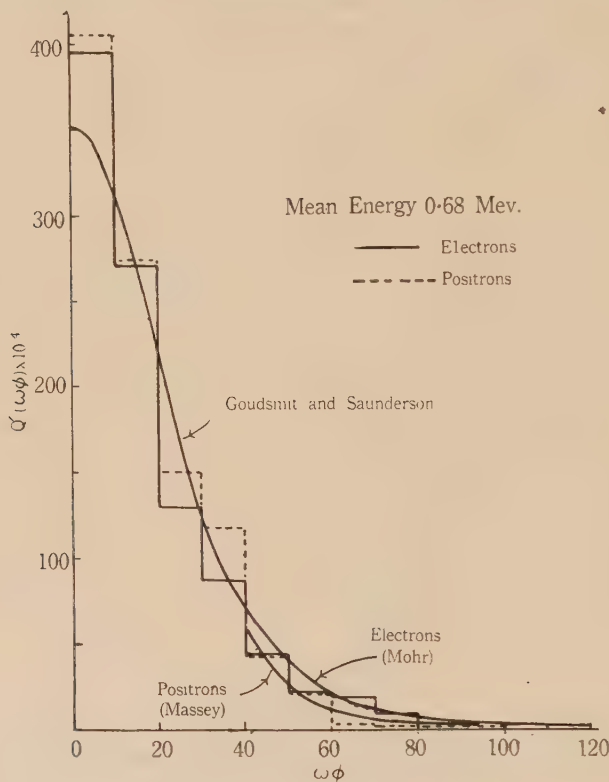


Fig. 3.—Angular distribution of scattering in the range 0.56–0.82 Mev.

(ii) *Scattering Probability*.—The distributions are most conveniently expressed as a function  $Q'(\omega\phi)$  of  $(\omega\phi)$  where

$$Q'(\omega\phi) = \frac{1}{S} Q(\omega\phi)$$

and  $Q(\omega\phi)$  is the probability that a particle incident on the foil will be scattered with the product  $\omega\phi$  lying between  $\omega\phi$  and  $\omega\phi + d\omega\phi$ , and at the same time possess an emergent branch with a projected length exceeding  $\lambda_c = 2$  cm.  $\omega$  is the total energy of the particle expressed in  $mc^2$  units, and  $\phi$  the projected angle in degrees.

The function  $Q'(\omega\phi)$  changes only slowly with  $\omega$ , so that three groupings are sufficient for the range of energies 0.39-1.1 Mev. These are shown together with the experimental data in Figures 2, 3, and 4.

(iii) *Small Angle Scattering*.—A thin foil was chosen so as to give an approach to single scattering at the larger angles, but the lower limit to the thickness was set by the presence of gas scattering. The mean angle of scattering in the foil used is equivalent to the mean deflection of a particle in a length of 24 cm. in

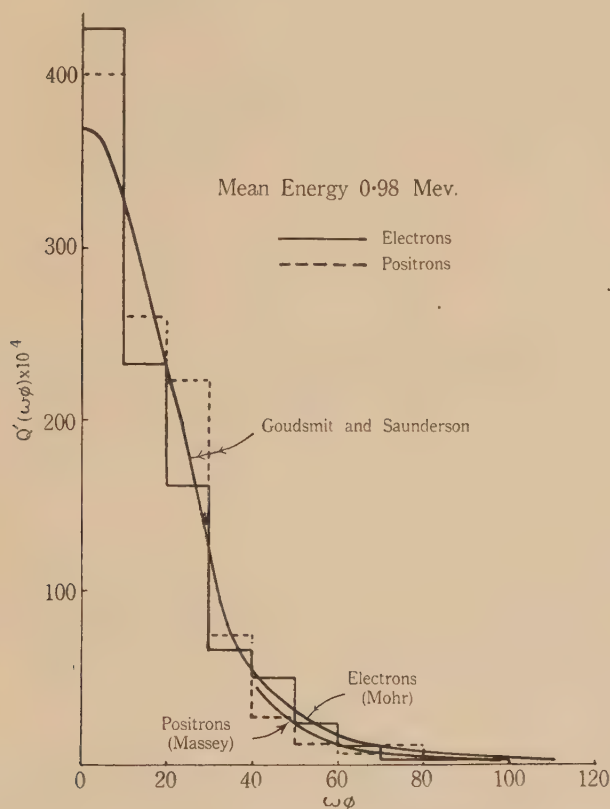


Fig. 4.—Angular distribution of scattering in the range 0.82-1.2 Mev.

the gas of the chamber. The effects of gas scattering are most important for small angle deflections near the foil, the exact contribution being difficult to estimate accurately, but Williams' formula indicates an effective increase in the effective mean angle of about 15 per cent. The variation of the experimental values of the mean angle of scattering with energy is shown in Figure 5. The results for electrons fall between the values of  $\alpha_{AV}$  calculated from the Goudsmit and Saunderson formula for  $\omega\phi < 100$   $mc^2$  degree units, and with the  $\alpha_m$  of Williams' theory.



The mean angle for electrons appears to be slightly greater than that for positrons throughout the range, although the difference is of the same order as the statistical errors of the experiment. A closer comparison with Williams'

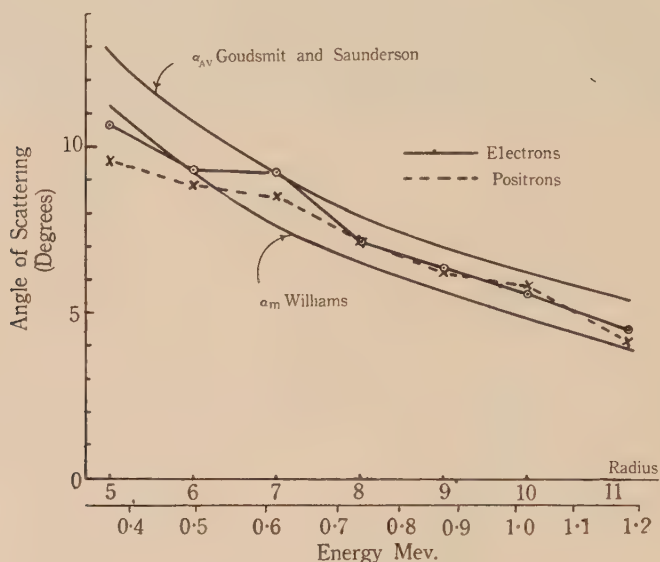


Fig. 5.—Mean angle of scattering as a function of particle energy.

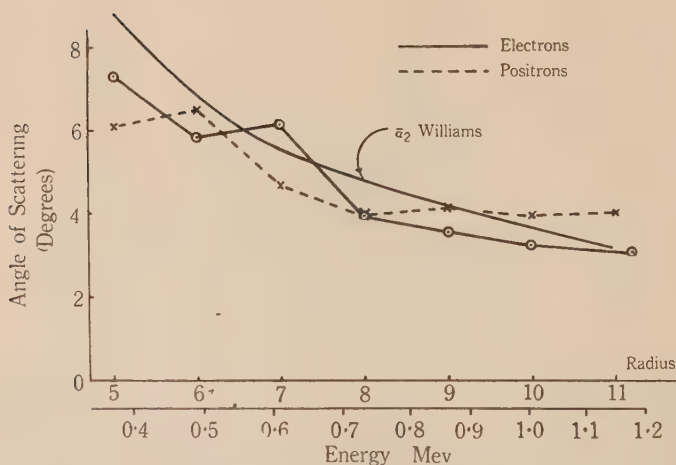


Fig. 6.—Mean multiple angle of scattering as a function of particle energy.

theory of multiple scattering is afforded by the calculation of his mean projected angle  $\bar{\alpha}_2$ , after excluding angles due to a single scattering. These results are shown in Figure 6 and indicate a slightly better agreement.

(iv) *Large Angle Scattering*.—Here more significant evidence of the difference between positron and electron may be obtained. It is possible to make comparisons for  $\omega\varphi$  greater than about  $50 mc^2$  degree units, where the single scattering effect is not greatly diluted by the multiple scattering process. An upper limit to  $\varphi$  is set by the difficulty of observing deflections of greater than about  $50^\circ$ .

TABLE I  
PROJECTED SINGLE SCATTERING IN GOLD FOIL

	Radius of Curvature (cm.)	Mean Energy (Mev.)	Angular Range ( $\varphi_1, \varphi_2$ )	Total No. of Particles	No. of Scattered Particles	Scattering Probability $N'$ ( $\varphi_1, \varphi_2$ ) Expt. Theory (in units of $10^{-2}$ )	
Positrons	5 and 6	0.45	$27^\circ$ – $54^\circ$	301	12	4.0	3.9
	7 and 8	0.68	$21^\circ$ – $43^\circ$	376	13	3.5	3.5
	9, 10, and 11	0.95	$17^\circ$ – $33^\circ$	189	5	2.6	2.8
Electrons	5 and 6	0.45	$27^\circ$ – $54^\circ$	338	22	6.5	8.8
	7 and 8	0.68	$21^\circ$ – $43^\circ$	318	18	5.7	6.4
	9, 10, 11, and 12	0.99	$17^\circ$ – $33^\circ$	304	13	4.3	4.7

In Table 1,  $N'(\varphi_1, \varphi_2)$  indicates the probability that an emergent track having a length greater than 2 cm. shall be a scattering event for which the product  $\omega\varphi$  lies between 50 and  $100 mc^2$  degree units. Columns 3 and 4 indicate the corresponding angular limits for a particle of the mean energy in each range.

## VI. DISCUSSION

(i) So far as the numbers of scattering events permit comparison to be made, the distinctive theoretical scattering probabilities for positrons and electrons appear to be approximately correct in each of the three groups covering the range 0.39–0.97 Mev.

(ii) We observe that the absolute scattering cross section per atom is given by

$$\Phi(\theta)d\Omega = g(Z, \beta) \cdot F(\theta) \cdot d\Omega \text{ (cm.}^2 \text{ units)}$$

for a nuclear encounter which scatters the particles at an angle  $\theta$  into a solid angle  $d\Omega$ . The projection method only involves the scattering cross section implicitly. If the "single scattering tail" of projected deflections greater than an angle  $\varphi_1$  is in agreement with the theoretical form under projection, it indicates that the absolute scattering cross section  $\Phi(\theta)d\Omega$  is also approximately correct for space scattering angles in excess of  $\varphi_1$ . In the light of (i) and in view of the fact that the large angle scatterings in Figures 2, 3, and 4 do follow the theoretical

form, we believe that the absolute scattering intensity is in approximate agreement with theory.

(iii) In Table 2 the ratio of electron to positron scattering probability ( $e^-/e^+$ ) is seen to be consistent with the theoretical result. The statistical significance of the experimental data may be determined with the application of Fisher's(11) independent probability test. For each energy range a  $2 \times 2$

TABLE 2  
RATIO OF ELECTRON TO POSITRON SCATTERING AT LARGE ANGLES

Mean Energy (Mev.)	Angular Range ( $\phi_1, \phi_2$ )	Ratio of Scattering Probability	
		Expt.	Theory
0.45	27°-54°	1.6	2.3
0.68	21°-43°	1.6	1.8
0.97	17°-33°	1.6	1.7

table is formed from the number of scattered electrons and positrons observed, and each corresponding total. The test yields a value of  $\chi^2=12.0$  with 6 degrees of freedom giving a chance probability of 0.065.

This result is just outside the 0.05 level of statistical significance, but it indicates that the spin orbital interaction must be taken into account in describing the scattering process.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Atomic Physics Section, C.S.I.R., University of Melbourne.

The author is indebted to Dr. C. B. O. Mohr, Senior Lecturer, Physics Department, University of Melbourne, for his criticism and help in the preparation of the manuscript, and to Mr. J. A. McDonell, Physics Department, for his assistance in making the calculations.

## VIII. REFERENCES

- (1) FOWLER, W. A., and OPPENHEIMER, JAQUELINE.—*Phys. Rev.* **54**: 320 (1938).
- (2) WILLIAMS, E. J.—*Proc. Roy. Soc. A* **169**: 531 (1938).
- (3) GOUDSMIT, S., and SAUNDERS, J. L.—*Phys. Rev.* **57**: 24 (1940); **58**: 36 (1940).
- (4) BARTLETT, J. H., and WATSON, R. E.—*Proc. Amer. Acad. Arts Sci.* **74**: 53 (1940).
- (5) MASSEY, H. S. W.—*Proc. Roy. Soc. A* **181**: 14 (1942).
- (6) Ó'CEALLAIGH, C., and MACCÁRTHAIGH, M. D.—*Proc. Roy. Irish Acad.* **A 50**: 13 (1944).
- (7) TALLQVIST, HJ.—*Acta Soc. Sci. Fenn.* (Comm. Phys.-math.) **6** (3): (1933).
- (8) MOHR, C. B. O.—*Proc. Roy. Soc. A* **182**: 189 (1943).
- (9) BARKER, F. C.—*J. Sci. Instrum.* **25**: 165 (1948).
- (10) CHASE, C. T., and COX, R. T.—*Phys. Rev.* **58**: 243 (1940).
- (11) FISHER, R. A.—“Statistical Methods for Research Workers”, 8th Ed., p. 97. (Oliver & Boyd: London, 1941.)

## APPENDIX

By J. A. McDONELL\*

*Projection of the Scattering Laws*

Ó'Ceallaigh and MacCárthaigh(6) have shown that when the general scattering law is written in the form (1), the projected distribution is given by

$$P(\varphi) = \frac{Ntg(Z, \beta)}{a} \sec^2 \varphi \int \int \frac{F(\theta) \sin \theta}{\sqrt{\sec^2 \theta - \sec^2 \varphi}} d\theta dy,$$

where the integral is to be evaluated over the  $(\theta, y)$  regions determined by

$$\begin{aligned} \cos \theta &= \lambda \cos \varphi / \sqrt{\lambda^2 + (a \pm y)^2} \\ -a &\leq y \leq a \\ \lambda &\geq \lambda_c. \end{aligned}$$

If we perform the  $\theta$ -integration first, we obtain

$$P(\varphi) = \frac{2Ntg(Z, \beta)}{a} \sec^2 \varphi \int_{-a}^a \int_{\varphi}^{\theta_M} \frac{F(\theta) \sin \theta}{\sqrt{(\sec^2 \theta - \sec^2 \varphi)}} d\theta dy$$

$$\text{with } \cos \theta_M = \lambda_c \cos \varphi / \sqrt{\lambda_c^2 + (a + y)^2},$$

which is the form given by Ó'Ceallaigh and MacCárthaigh.

However, if we perform the  $y$ -integration first, we obtain

$$P(\varphi) = \frac{2Ntg(Z, \beta)}{a} \sec^2 \varphi \int_{\varphi}^{\theta_L} \int_{-a}^{A(\varphi)} \frac{F(\theta) \sin \theta}{\sqrt{(\sec^2 \theta - \sec^2 \varphi)}} dy d\theta$$

$$\text{where } A(\varphi) = a - \lambda_c \sqrt{\cos^2 \varphi \sec^2 \theta - 1}, \quad \cos \theta_L = \frac{\rho}{\sqrt{\rho^2 + 1}} \cos \varphi,$$

$$\text{and } \rho = \lambda_c / 2a$$

$$\begin{aligned} \text{i.e. } P(\varphi) &= 4Ntg(Z, \beta) \sec^2 \varphi \int_{\varphi}^{\theta_L} \frac{F(\theta) \sin \theta}{\sqrt{(\sec^2 \theta - \sec^2 \varphi)}} \\ &\quad \times (1 - \rho \sqrt{\cos^2 \varphi \sec^2 \theta - 1}) d\theta. \end{aligned}$$

The substitution

$$\cos \theta = \cos \varphi \cos \chi$$

reduces this integral to the form

$$P(\varphi) = 4Ntg(Z, \beta) \int_0^{\arccot \rho} F(\theta) [\cos \chi - \rho \sin \chi] d\chi \dots \dots (2)$$

\* Physics Department, University of Melbourne.

From the Goudsmit and Saunderson theory we calculated values of  $f(\theta) = Ntg(Z, \beta)F(\theta)$  at  $5^\circ$  intervals of  $\theta$  from  $0$  to  $60^\circ$ , and using these,  $P(\varphi)$  was calculated numerically from (2), for our case of  $\rho=1$ . Since  $f(\theta)$  is normalized so that

$$\int_0^\pi 2\pi f(\theta) \sin\theta \, d\theta = 1,$$

the ratio  $S$  as previously defined, will be given by

$$S = \int_0^\pi P(\varphi) d\varphi.$$

The calculated values of  $S$  are

Energy (Mev.)	0.45	0.7	0.9
$S$	0.78	0.83	0.88

Finally,  $Q'(\omega\varphi)$  was calculated, from

$$Q'(\omega\varphi) = \frac{\pi}{180\omega} \cdot \frac{1}{S} P(\varphi).$$



# COSMIC RAY MEASUREMENTS MADE DURING THE CRUISE OF H.M.A.S. WYATT EARP 1947-48

By D. E. CARO,\* P. G. LAW,\* and H. D. RATHGEBER\*

[Manuscript received July 2, 1948]

## Summary

This paper describes the cosmic ray equipment used and the results obtained on the cruise of the *Wyatt Earp* between Melbourne, Australia, and the Antarctic Continent during the early part of 1948. Measurements of the total cosmic radiation over a narrow solid angle, the meson content over the same angle, and the total content over the full hemisphere were made. The difficulty of carrying out multiple correlations, when the independent variables are closely related, is discussed. The final correlation results show barometric coefficients of  $-0.144$  per cent. per millibar (mb.) for the narrow angle total radiation,  $-0.114$  for the meson content, and  $-0.192$  for the total wide angle record. Long term temperature coefficients of  $-0.063$  per cent. per degree C.,  $-0.051$  per cent. per degree C., and  $-0.012$  per cent. per degree C., respectively, are obtained. The need to separate the daily and long-term temperature coefficients in regions where continental climates exist is shown. Daily temperature coefficients of  $0.14$  per cent. per degree C.,  $0.24$  per cent. per degree C., and  $0.16$  per cent. per degree C. are obtained. It is found that no measurable latitude effect remains after pressure and temperature corrections have been made. The ratios of the total to meson content and the narrow angle intensity to wide angle intensity are independent of latitude south of Melbourne. A possible effect due to a magnetic disturbance on February 15 is noted, but later more severe storms did not produce any noticeable result. The results obtained are compared with those of other workers.

## I. INTRODUCTION

The research programme for the Australian National Antarctic Research Expedition of 1947-48 included plans for cosmic ray work at three stations. Physicists with cosmic ray apparatus were landed at Heard and Macquarie Islands to make observations throughout the year 1948, and apparatus was also installed on board H.M.A.S. *Wyatt Earp* to record cosmic ray intensities during her cruise in Antarctic waters early in the same year.

The following account describes the equipment used on board the *Wyatt Earp*, and the results obtained between February 8 and April 5, 1948. It is hoped to publish reports on the island activities during 1949, and a complete account of the whole programme of cosmic ray research will finally appear in the collected papers of the A.N.A.R.E.

Figure 1 shows the route followed by the *Wyatt Earp*. From the chart it is seen that the cruise covered a latitude range from  $38^{\circ}$  S. to  $67^{\circ}$  S.†

\* Physics Department, University of Melbourne.

† One author (P.G.L.) accompanied the ship and operated the equipment. He was assisted by Mr. E. McCarthy, the Expedition's geophysicist on board the *Wyatt Earp*.

## II. DESCRIPTION OF EQUIPMENT

The cosmic ray recorder, with which the results given in this paper were obtained, consisted basically of six Geiger counter trays as shown in Figure 2. Each tray contained six counters connected in parallel. Electronic circuits record threefold coincidences between trays 1, 2, and 3, thus recording the total intensity of cosmic rays over a comparatively narrow solid angle directed vertically. Threefold coincidences between trays 2, 3, and 4 record the hard component of the cosmic radiation, the soft component (mostly electrons) being



Fig. 1.—Track chart of H.M.A.S. *Wyatt Earp*.

incapable of penetrating the 10 cm. of lead between trays 3 and 4. The solid angle subtended by this counter telescope is the same as that for the total content from trays 1, 2, and 3. Twofold coincidences between trays 5 and 6 record the total radiation incident over practically the full hemisphere.

Figure 3 shows the circuit diagram of the recorder equipment. Geiger counter pulses from trays 1, 2, and 3 are applied to the grids of V1, V2, and V3 respectively. These three valves operate as a threefold Rossi coincidence circuit, producing a large positive pulse across their plate load whenever the pulses on their grids are coincident. This positive pulse is applied to the grid of V4. V4 and V5 are connected as a single shot multivibrator, V5 being normally non-conducting and V4 normally conducting. The screen of V5 is used as an auxiliary plate for the multivibrator circuit. When a positive pulse from V1,

V2, and V3 arrives at the grid of V4, grid current flows and C4 is charged. At the end of the pulse the grid of V4 is consequently carried negative. This negative pulse trips the single shot multivibrator and actuates the mechanical register K1. This register then records the threefold coincidences from trays 1, 2, and 3. The four decade telephone register used is of Swiss manufacture. In this circuit it is capable of count rates up to 25 counts per second.

Valves V6 to V10 inclusive are connected in an identical circuit and the register K2 records threefold coincidences from trays 2, 3, and 4.

A twofold Rossi coincidence circuit, V11 and V12, detects coincidences from trays 5 and 6. This circuit is followed by a discriminator V13 which passes short negative pulses to a scale-of-two formed by V14 and V15. This scale is a normal pentode type. Pentagrid tubes were used because suitable battery

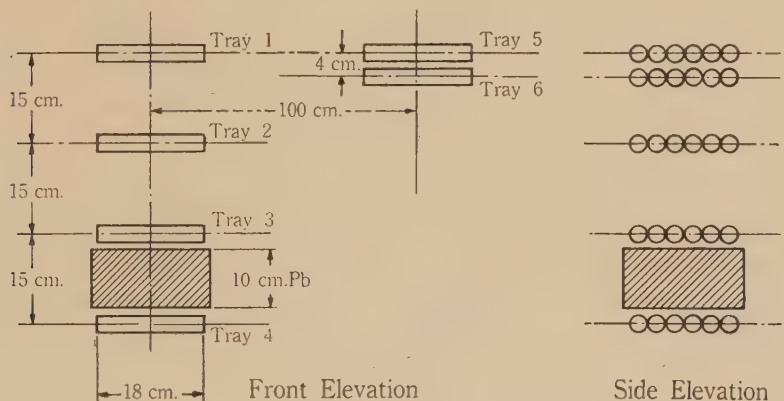


Fig. 2.—Geiger counter layout

pentodes without internal suppressor-cathode connection were not available. The control VR1 enables the scaling circuit to be balanced to compensate for any small asymmetry. The first scale is followed by a buffer amplifier V16 and a second scale V17, V18. This results in a total scale factor of four. The output of the final scale operates the register K3 through the buffer amplifier V19 and power amplifier V20. The resolving time of the coincidence circuits is of the order of  $20 \mu$  sec., resulting in an upper limit to the number of accidentals for the twofold circuit of about 20 per hour. The threefold circuits will have less than one-half this number of accidentals. Each valve has its own individual filament-dropping resistor, the filaments being operated from the +2 volt line. Figure 4 shows the circuit of the power unit used to operate the recorder. Primary power is obtained from 12-volt lead-acid accumulators which were charged from the ship's D.C. supply. The battery was connected to the system earth at 2 volts from its positive end. The filaments were then operated between this point and +2 volts and the vibrator between earth and -10 volts on the battery.

The vibrator supplies power to a special three-winding transformer. The centre-tapped 220-0-220 volt secondary winding is connected back to the vibrator to produce +220 volts by synchronous rectification. After filtering, this supply is stabilized at +150 volts by V23. A second, 110-volt winding produces -75 volts stabilized, V22 acting as a half-wave rectifier and V24 as a stabilizer. The third, 1300-volt winding produces the Geiger counter high tension. V21 acts as a half-wave rectifier, producing about 950 volts across the 17 voltage stabilizer neons. SW2 and VR3 provide coarse and fine adjustments for the counter high tension. The voltage at the counter cathodes was adjusted to -850 volts. Their anodes were returned to +150 volts (see Fig. 3) resulting in a total counter potential of 1000 volts.

The Geiger counters used were 3.1 to 3.3 cm. in diameter and had an effective counting length of 18 cm. The cathode was prepared by evaporating a copper film on to the inside of the glass envelope. The counters were filled to a pressure of 7 mm. of ether and the pressure of argon was adjusted to give a starting voltage of  $900 \pm 10$  volts. This corresponded to an argon pressure of about 7 cm. Only counters with plateaux exceeding 200 volts were accepted.

Although the laboratory was kept comfortable by an electric heater, no special precautions were taken to keep the temperature constant. As the dew point of the ether filling is approximately  $-50^{\circ}\text{C.}$ , the counters operated satisfactorily at all temperatures encountered(1). Early in the work it was found that vibration from the ship's engines caused cracks in the seals of the Geiger counter tubes. This defect was remedied by mounting the entire equipment on one inch of sponge rubber. Excessive wear resulted in the breakdown of the registers after a few hundred thousand counts. In future equipment it is proposed to introduce sufficient scale factors in the electronic circuits to keep the count rate at the registers below 1000 per hour. The registers were photographed at hourly intervals with a 16 mm. camera timed by a ship's chronometer. The apparatus was placed in a small laboratory on the top deck. The roof, which was only about 25 cm. above the top of the apparatus, consisted of about 5 g./cm.<sup>2</sup> wood and 4 g./cm.<sup>2</sup> mild steel. The apparatus was surrounded by a box with aluminium walls of 1 g./cm.<sup>2</sup>.

### III. RESULTS AND DISCUSSION

A nearly continuous record was obtained from all three counter sets. The interruption in all three records, which began on February 19, was caused by violent rolling of the ship of more than  $45^{\circ}$  each side. The whole apparatus was torn from its base and was badly damaged. The other interruptions were due mainly to the failures of the telephone registers. Registers of the same make obtained before the war showed a much longer life.

It was intended in planning the experiment to compare the cosmic ray records with the height of several pressure levels obtained by simultaneous radio-sonde flights. This would have permitted correction for the atmospheric density effect(2) caused by the decay of the mesons, and the detection of any geomagnetic latitude effect existing at latitudes above the "knee"(3). The use



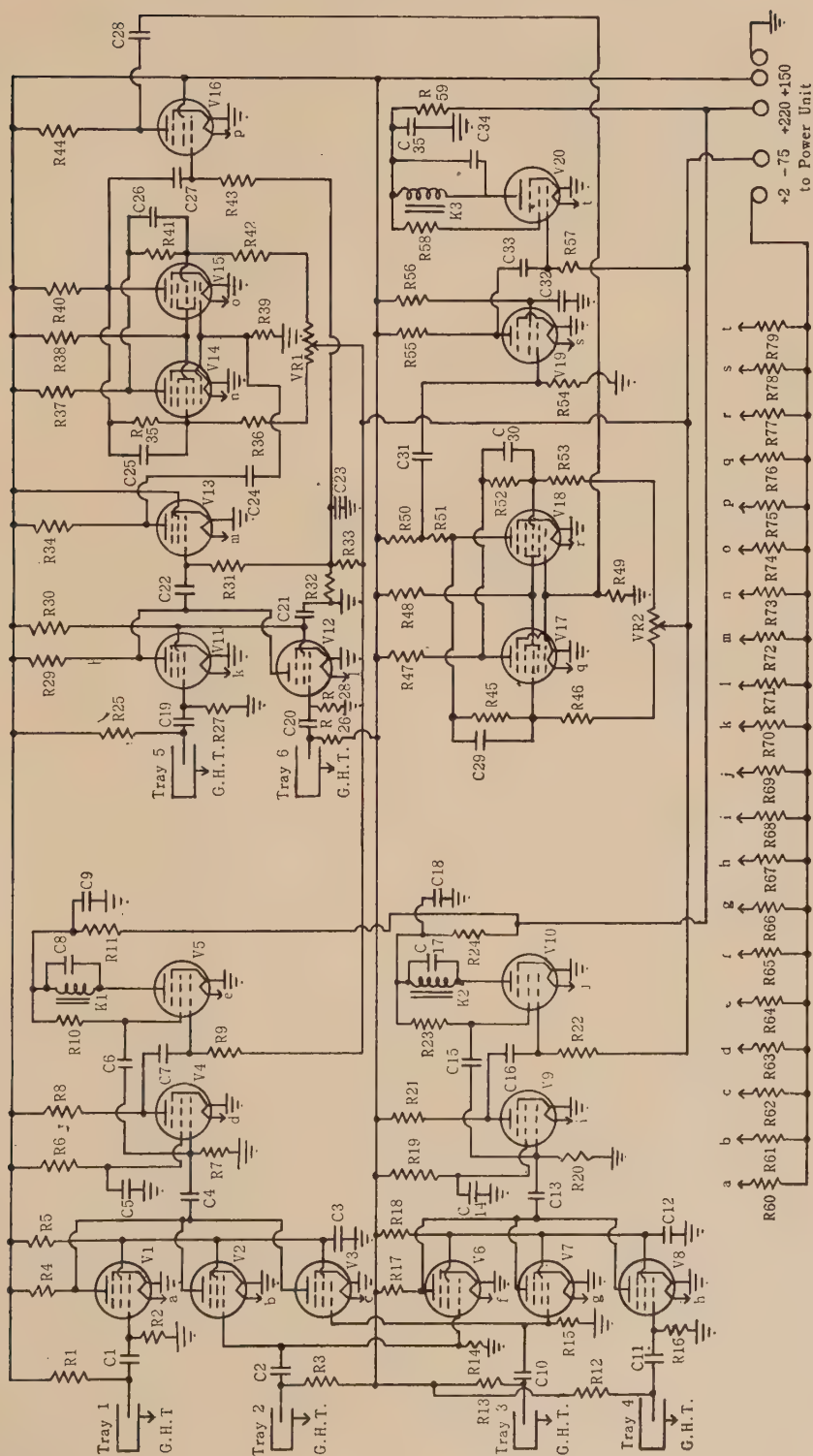


Fig. 3.—Cosmic ray recorder : schematic circuit diagram.



## PARTS LIST FOR CIRCUITS (FIGS. 3 AND 4)

## Valves

V1, V2, V3, V4, V6, V7, V8, V9, V11, V12, V13,	
V16, V19 .. .. .	1L4
V5, V10, V20, V22 .. .. .	1S4
V14, V15, V17, V18 .. .. .	1R5
V21 .. .. .	1Z2
V23 .. .. .	VR150/30
V24 .. .. .	VR75/30

## Resistors

R1, R3, R12, R13, R25, R26 .. .. .	50-megohm
R85, R86, R87 .. .. .	5-megohm
R54 .. .. .	2-megohm
R57, R82 .. .. .	1-megohm
R31, R43 .. .. .	500,000-ohm
R35, R41, R45, R52 .. .. .	250,000-ohm
R2, R4, R7, R8, R9, R14, R15, R16, R17, R20,	
R21, R22, R27, R28, R29, R33, R36, R37,	
R40, R42, R46, R47, R53, R55 .. .. .	200,000-ohm
R51 .. .. .	150,000-ohm
R32 .. .. .	100,000-ohm
R6, R19, R38, R39, R48, R49, R50, R56 .. .. .	50,000-ohm
R5, R18, R30 .. .. .	30,000-ohm
R34, R44 .. .. .	20,000-ohm
R10, R23, R58 .. .. .	10,000-ohm
R83 .. .. .	5,000-ohm
R84 .. .. .	2,500-ohm
R11, R24, R59 .. .. .	2,000-ohm
R60, R61, R62, R63, R65, R66, R67, R68, R70,	
R71, R72, R73, R74, R75, R76, R77, R78 .. .. .	8-ohm
R64, R69, R79, R81 .. .. .	4-ohm
R80 .. .. .	2-ohm
VR1, VR2, VR3 .. .. .	100,000-ohm potentiometers

## Condensers

C39 .. .. .	12-microfarad
C9, C18, C35 .. .. .	8-microfarad
C38, C40, C41 .. .. .	4-microfarad
C42 .. .. .	2-microfarad
C8, C17, C34 .. .. .	0.5-microfarad
C3, C5, C6, C7, C12, C14, C15, C16, C21, C23, C32,	
C33, C43, C44, C45, C46, C47, C48 .. .. .	0.1-microfarad
C31 .. .. .	0.02-microfarad
C36, C37 .. .. .	0.002-microfarad
C25, C26, C29, C30 .. .. .	0.001-microfarad
C4, C13 .. .. .	200-microfarad
C1, C2, C10, C11, C19, C20 .. .. .	100-microfarad
C22, C24, C27, C28 .. .. .	50-microfarad

of an arbitrary sea-level temperature coefficient(4, 5) chosen to make any latitude effect disappear is open to question. Unfortunately no more space on the ship was available for using the radio-sonde equipment.

(i) *Statistical Analysis*.—In the absence of upper-air data, it was decided to try to eliminate the meteorological factors by multiple correlation calculation of cosmic ray intensity with barometric pressure, external air temperature, and geographic latitude. The measurements were subdivided into three groups: group 1 from latitude 38° S. to 51° S., group 2 from 51° S. to 65° S., and group 3

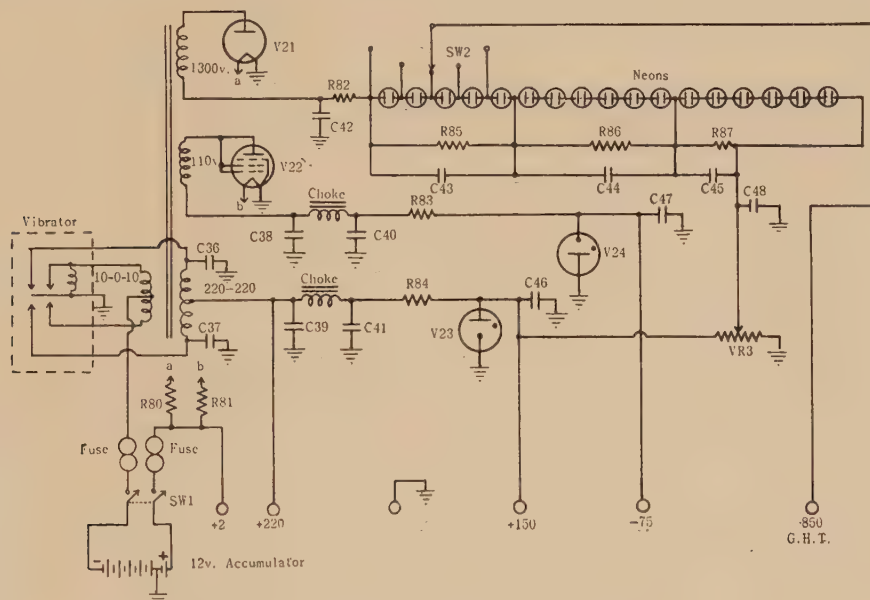


Fig. 4.—Cosmic ray recorder: power unit.

from 65° S. to 67° S. As the latitude of the last group was nearly constant, the latitude was dropped from the correlation calculation. The purpose of the subdivision was to find whether any significant variation of the regression coefficients with latitude existed. To reduce the influence of the random variations, four-hourly means were used. No noticeable amount of statistical information was lost in this way; barometric pressure and air temperature varied only slightly during such intervals (see Fig. 6).

Calculating machines were used to carry out the statistical analysis by known methods(6). The variations in each group of the means and standard deviations of pressure, temperature, and latitude, as well as of the simple correlation coefficients between pairs of these variables, are due to the gaps in some of the cosmic ray measurements.

Table 1 contains the results of these calculations. A comparison of the standard deviations of the cosmic ray intensities  $\bar{\sigma}_C$  with the corresponding

TABLE 1  
STATISTICAL ANALYSIS OF COSMIC RAY MEASUREMENTS: "WYATT EARP" CRUISE 1947-48

Statistical Values	Symbol	Group 1: Lat. 38° S. to 51° S.				Group 2: Lat. 51° S. to 65° S.				Group 3: Lat. 65° S. to 67° S.		
		Count 1*	Count 2*	Count 3*		Count 1	Count 2	Count 3		Count 1	Count 2	Count 3
1. Mean latitude (° S.)	$\bar{L}$	42.3	42.1	42.3		57.7	56.8	56.9		66.4	66.4	66.4
2. Mean pressure (mb.)	$\bar{P}$	1013.3	1014.1	1013.3		990.6	989.3	989.4		984.2	984.4	984.2
3. Mean temperature (° C.)	$\bar{T}$	13.8	13.9	13.8		3.8	3.7	3.7		-2.7	-2.7	-2.6
4. Mean cosmic ray intensity (counts/hour)	$\bar{C}$	3997	2867	19356		4158	2976	20166		4281	3049	20683
5. Number of recordings	$N$	95	91	95		50	83	87		126	133	151
6. Standard deviation of latitude	$\sigma_L$	4.60	4.56	4.60		4.52	3.81	3.75		7.80	7.81	7.44
7. Standard deviation of pressure	$\sigma_P$	12.86	12.46	12.85		11.60	10.79	10.58		2.19	2.25	2.10
8. Standard deviation of temperature	$\sigma_T$	3.10	3.12	3.10		2.89	2.45	2.39		82.7	56.5	337.4
9. Standard deviation of cosmic ray intensity	$\sigma_C$	101.2	70.3	500.3		138.0	75.5	480.1				
<i>Simple Correlation Coefficients</i>												
10. Cosmic ray-pressure	$r_{CP}$	-0.77	-0.64	-0.97		-0.79	-0.58	-0.94		-0.57	-0.47	-0.87
11. Cosmic ray-temperature	$r_{CT}$	-0.32	-0.41	-0.57		-0.62	-0.23	-0.55		-0.22	-0.18	-0.01
12. Cosmic ray-latitude	$r_{CL}$	+0.51	+0.47	+0.87		+0.72	+0.22	+0.52		-0.02	-0.05	-0.05
13. Pressure-temperature	$r_{PT}$	+0.61	+0.59	+0.61		+0.80	+0.59	+0.59				
14. Pressure-latitude	$r_{PL}$	-0.88	-0.88	-0.88		-0.82	-0.61	-0.60				
15. Temperature-latitude	$r_{TL}$	-0.76	-0.75	-0.76		-0.96	-0.93	-0.93				
<i>Partial Regression Coefficients</i>												
16. Cosmic ray-pressure (% of mean/mb.)	$b_{CP}$	-0.285 ± 0.024	-0.211 ± 0.033	-0.184 ± 0.011		-0.188 ± 0.039	-0.166 ± 0.027	-0.216 ± 0.010		-0.144 ± 0.018	-0.114 ± 0.018	-0.192 ± 0.009
17. Cosmic ray-temperature (% of mean/° C.)	$b_{CT}$	-0.016 ± 0.022	-0.071 ± 0.029	+0.017 ± 0.010		+0.342 ± 0.095	-0.009 ± 0.081	-0.073 ± 0.031		-0.063 ± 0.019	-0.051 ± 0.019	-0.012 ± 0.009
18. Cosmic ray-latitude (% of mean/° lat.)	$b_{CL}$	-0.450 ± 0.083	-0.371 ± 0.110	+0.062 ± 0.037		+0.812 ± 0.208	-0.155 ± 0.172	-0.175 ± 0.064				
19. Coefficient of multiple correlation	$\bar{R}$	0.84	0.68	0.94		0.84	0.57	0.94		0.61	0.50	0.87
20. Minimum probable correlation	$R_{min}$	0.77	0.56	0.92		0.74	0.42	0.92		0.49	0.35	0.82
21. Standard error of estimate (counts/hour)	$\bar{s}$	55.5	52.2	120		76.5	62.3	165		65.1	49.0	165
22. Random error $\frac{1}{\sqrt{C}}$ (counts/hour)	$\sigma_C$	31.6	26.8	70		32.2	27.3	71		32.7	27.6	72
23. Maximum probable correlation	$R_{max}$	0.95	0.92	0.99		0.97	0.93	0.99		0.91	0.87	0.97

\* Count 1: Total intensity, narrow angle. Count 2: Hard component, narrow angle. Count 3: Total intensity, wide angle.

standard errors of estimate  $\bar{S}$ , which are the calculated standard deviations of the corrected intensities, shows that a considerable reduction has been achieved. The random nature of cosmic rays itself would, even in the absence of any other factors, cause a statistical distribution of the four-hourly means. This random error, the standard deviation of a Poisson distribution, is equal to  $\frac{1}{2}\sqrt{\bar{C}}$ .<sup>\*</sup> The fact that this random error  $\bar{\sigma}_C$  is in all cases only about half of the standard error of estimate shows that a considerable amount of unaccountable fluctuation remains. Another way to consider the problem is to calculate the maximum possible correlation when only these random errors are taken into account. Consider the coefficient of multiple correlation  $\bar{R}$ . The square of this quantity represents the fraction of the variance (square of the standard deviation) which can be attributed to the independent variables of the correlation (7). Thus a fraction  $(1-\bar{R}^2)$  of the variance is unrelated to the independent variables. If we assume that this independent fraction is the random error we obtain

$$(1-\bar{R}_{max}^2) = \bar{\sigma}_C^2 / \sigma_C^2 \quad \text{or} \quad \bar{R}_{max} = (1 - \bar{\sigma}_C^2 / \sigma_C^2)^{\frac{1}{2}}$$

where  $\bar{R}_{max}$  is the maximum possible correlation. The coefficients of multiple correlation are all smaller than this maximum. This confirms the above conclusion that a considerable proportion, but not all causes of fluctuations, has been accounted for in these calculations.

A survey of the regression coefficients shows great variation between the different groups, e.g. the cosmic ray latitude coefficient for count 1, which is  $-0.450 \pm 0.083$  in group 1, changes to  $+0.812 \pm 0.208$  in group 2, all errors being standard errors. Despite these variations, a time graph of the cosmic ray intensities, corrected with these coefficients in their respective groups, yielded an essentially flat curve with only normal variations about the mean. The explanation of this is found in the very high correlation coefficients of pressure-latitude  $r_{PL}$  and temperature-latitude  $r_{TL}$  and the slightly lower one of pressure-temperature  $r_{PT}$ . If the correlation coefficients were 1, the variables could be exchanged one against the other without affecting the results. In our case the corrections for the latitude change of the cosmic ray intensities split up into the regression coefficients for pressure, temperature, and latitude in a haphazard way which is determined by accidental features in the records. The smoothing out obtained by the application of these coefficients is internally consistent, showing that their variations cancel each other but that no physical significance can be attached to any of them.

(ii) *Reduction of Measurements to Standard Pressure and Temperature.*—If we consider group 3 the position changes. The observations were taken at nearly constant geographic latitude, and as there is no correlation between pressure and temperature, there can be no such splitting up as in groups 1 and 2. The regression coefficients of this group were taken as the ones having physical significance and all cosmic ray measurements were corrected with them to 1000 mb. and 0° C. The result shown in Figure 6 justifies the procedure. The

<sup>\*</sup> The factor  $\frac{1}{2}$  is due to averaging over four-hourly intervals.



variations due to barometric pressure are evidently smoothed out. The temperature corrections, which are small except at the lowest latitudes, straighten out the remaining trends between the highest and lowest latitudes.

(iii) *Latitude Effect*.—In these measurements the latitude effect remaining after barometric correction disappears on application of a temperature coefficient which is *not* taken arbitrarily, so as to make that latitude effect zero, but is derived from measurements at constant latitude. If it is assumed that the temperature effect is the same at all latitudes of the present experiment it can be stated that the latitude effect for the total wide angle radiation, for the total narrow angle radiation, and for the hard (10 cm. lead filter) narrow angle radiation does not exceed 0.5, 1, and 1 per cent. respectively. The fact that there is no measurable geomagnetic effect is confirmed by observing the curves between Commonwealth Bay, which is the nearest position to the south magnetic pole, and the Balleny Islands, which lie at the same geographic latitude, but the magnetic latitude of which is about 84° S.

(iv) *Barometric Effect*.—The fact that the pressure coefficient for the hard radiation is smaller than that for the total is explained by the filtering out of the low energy radiation(8). No significance need be attached to the difference in the temperature coefficients which, however, goes in the same direction.

The difference 0.48%/mb. between the pressure coefficients of count 1 and count 3 exceeds considerably the sum of their errors 0.03. The meson decay theory of the barometric effect provides the explanation. A change in height  $\Delta h$  of the meson-producing layers is a change  $\Delta h/\cos \alpha$  for rays forming an angle  $\alpha$  with the vertical. The absorbing mass increases in the same ratio  $1/\cos \alpha$ . Therefore the ratio of the pressure coefficients will be inversely proportional to the cosine or

$$\frac{0.144 \pm 0.018}{0.192 \pm 0.009} = \cos \alpha; \quad \alpha = 50^\circ \pm 10^\circ.$$

This is in satisfactory agreement with the average angle of incidence of the rays in the two counter arrangements. The possible difference in the energy spectrum is neglected.

Whilst many measurements of the barometric coefficient with ionization chambers have been reported(9) only a few measurements with Geiger counter coincidence apparatus have been described(10, 11). The former give on the average a barometric coefficient  $-0.13\%$ /mb. which compares well with our result,  $-0.19$  for count 3 (wide angle count), since a smaller coefficient is to be expected for the radiation measured in ionization chambers shielded by 10 cm. of lead.

A comparison with other coincidence measurements shows a difference in results which cannot be ascribed to experimental errors. Whilst Barnóthy and Forró(10) report a coefficient of  $-0.28\%$ /mb. for experimental conditions similar to our own count 1, we find only about half ( $-0.14$ ). Duperier(11) also finds  $-0.27\%$ /mb. for conditions corresponding to our count 3, which gives only  $-0.192$ .



We see no way in which the difference could be ascribed to a defect in our apparatus, which can be checked in two ways ; first by its absolute sensitivity, and secondly by the absorption in the 10 cm. of lead. The first is obtained from the following formula. The expected number of counts per hour  $C$  for unfiltered vertical radiation is approximately equal to  $3600IA^2/D^2$  in which  $A$  is the effective area of one tray,  $D$  the distance between extreme trays, and  $I$  the number of cosmic ray particles per unit area, second, and unit solid angle at 1013 mb. and  $0^\circ\text{C}$ . From the mean 3940, corrected to 1013 mb., we obtain

$$I = \frac{3940}{3600} \frac{30^2}{18^4} = 0.0093 \text{ particles cm.}^{-2} \text{ sec.}^{-1} \text{ steradian}^{-1}$$

which tallies well with precision determinations(12, 13), which give 0.010 and 0.0093. Our average absorption in 10 cm. of lead is

$$\frac{4010-2920}{4010} = 27 \pm 1\%,$$

which is higher than the best figure available(14) of  $22 \pm 1\%$ . This higher intensity of the soft component can be attributed to shower formation in the material of the roof. But even if it should be due to some instrumental defects, the difference is too small to account for the factor 2 in the value of the barometric effect.

(v) *Temperature Effect.*—It is even more difficult to compare our temperature coefficients with other results, which range from  $+1.6\%/^\circ\text{C}$ .(8) to  $-0.43\%/^\circ\text{C}$ .(15). Our results lie within these limits. In this connection we should like to draw attention to the fact that after the return to Melbourne, i.e. to a more continental climate, daily variations occurred which correlate with sea-level temperatures. Temperature coefficients of  $0.14 \pm 0.12\%/^\circ\text{C}$ . for count 1,  $0.24 \pm 0.08$  for count 2, and  $0.16 \pm 0.08$  for count 3, were obtained from scatter diagrams for all the days available in April. They are positive in contrast to the ones obtained from the more southerly measurements in a purely maritime climate. It is quite evident from the graph that the negative temperature coefficient from Table 1 as well as this positive one has to be applied to these days in April. The mean temperature for each day is determined and the cosmic ray readings are corrected accordingly with the negative (long term) temperature coefficient. These corrected readings are then further adjusted using the positive (daily) temperature coefficient to compensate for the individual temperature departures from their daily mean. The results of such a correction are shown in Figure 5. If we assume the meson theory to be correct we have to draw the conclusion that the daily variations of temperature affect the height of the meson-producing layers in a different way from long term variations of temperature. It seems necessary to separate both to obtain satisfactory temperature coefficients for the particular climate in which the measurements are made.

Other observations(15) comparable to our count 3, also give evidence of the same effect. A temperature coefficient  $-0.43 \pm 0.05\%/^\circ\text{C}$ . obtained from daily means is nearly compensated by the positive temperature coefficient of the daily temperature variation. The temperature coefficient  $-0.10 \pm 0.02\%/^\circ\text{C}$ .

obtained from hourly means is the sum of the long term coefficient and of that of the daily variation, which is equal to

$$0.43 \pm 0.05 - 0.10 \pm 0.02 = +0.35 \pm 0.07 \% / ^\circ \text{C.}$$

for the meteorological conditions at the time and place of the equipment.

(vi) *Magnetic and Other Disturbances.*—Other factors, such as magnetic storms(16), solar activity(17), and atmospheric density distribution are known to affect the cosmic ray intensity. It is hoped that a further correlation calculation with the available variables will further reduce the fluctuations not accounted for. Figure 6 shows only one correlated feature of cosmic ray intensity and

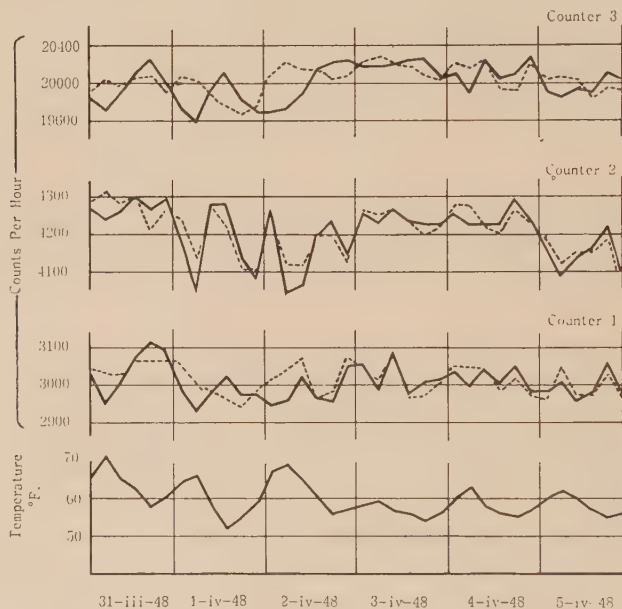


Fig. 5.—Daily variation of cosmic ray intensities with temperatures.

--- Corrected for daily mean temperatures.  
 — Corrected for four-hourly temperatures.

magnetic index (K sum). The magnetic disturbance of February 15 appears to be accompanied by a fall in cosmic ray intensity. Unfortunately the recordings during the fall in intensity were missed due to a failure of the apparatus. However, a definite record of the following rise extending over several days was made. No such effect is detectable during the much more intense magnetic storm of March 15. Comparison with other simultaneous measurements will show whether this effect is due to magnetic disturbance or to simultaneous barometric depression. The very pronounced spike on March 1, in count 3, is four times the standard error of estimate and thus can scarcely be explained as accidental. Study of the uncorrected curve and the barometric pressure curve reveals its cause. There is a well-defined positive random variation followed by a large negative one. At the rapid onset of the barometric depression the barometric correction is obviously too small. This effect has been observed before(18).

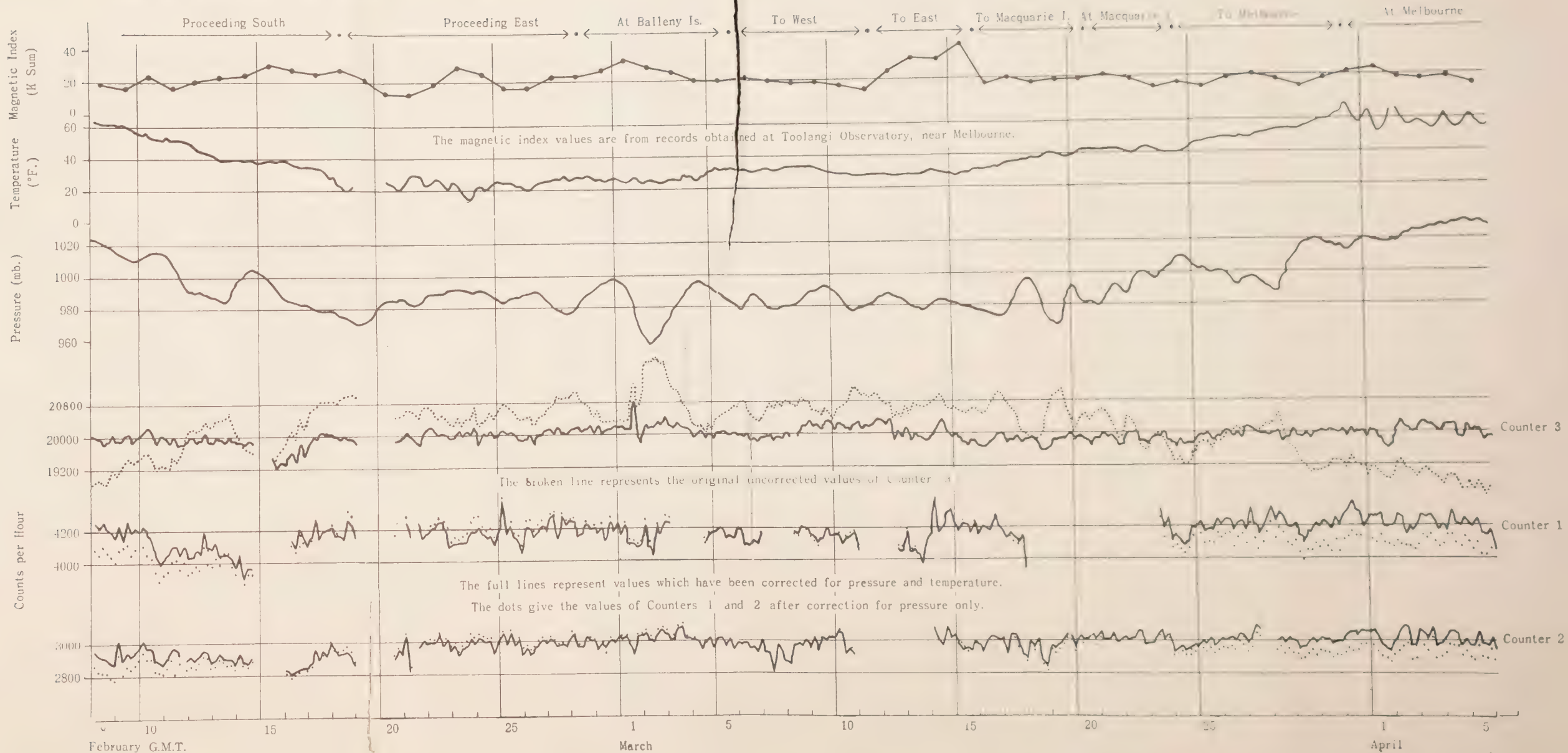


Fig. 6.—Record of four-hourly means of cosmic ray intensity, barometric pressure, temperature, and magnetic index. All times are G.M.T.

## IV. ACKNOWLEDGMENTS

The equipment was designed and built in the Physics Department, University of Melbourne, the cost being covered by a grant from the A.N.A.R. Expedition. The authors express their thanks to Professor L. H. Martin for his constant interest, and to Mr. E. McCarthy for his valuable help during the voyage. Mt. Stromlo Observatory and the Department of Mineral Resources have assisted by supplying sunspot, ionospheric, and magnetic data. The assistance of Misses D. B. Stoney and F. C. M. Stringer, who carried out much of the tedious correlation calculation, is acknowledged.

## V. REFERENCES

- (1) KORFF, S. A., SPATZ, W. D. B., and HILBERRY, N.—*Rev. Sci. Instrum.* **13**: 127 (1942).
- (2) RATHGEBER, H. D.—*Nature* **162**: 303 (1948).
- (3) JÁNOSSY, L.—“Cosmic Rays.” (Oxford, 1948.)
- (4) KORFF, S. A., and CLARKE, E. T.—*J. Franklin Inst.* **230**: 567 (1940).
- (5) GAST, P. F., and LOUGHRIDGE, D. H.—*Phys. Rev.* **58**: 194 (1940).
- (6) EZEKIEL, M.—“Methods of Correlation Analysis.” (New York, 1945.)
- (7) FISHER, R. A.—“Statistical Methods for Research Workers.” (London, 1936.)
- (8) MESSERSCHMIDT, W.—*Z. Phys.* **78**: 668 (1932).
- (9) DOAN, R. L.—*Phys. Rev.* **49**: 107 (1936).
- (10) BARNÓTHY, J., and FORRÓ, M.—*Z. Phys.* **100**: 742 (1936).
- (11) DUPIERIER, A.—*Terr. Magn. Atmos. Elect.* **49**: 1 (1944).
- (12) EHMERT, A., and TROST, A.—*Z. Phys.* **100**: 553 (1936).
- (13) GREISEN, K.—*Phys. Rev.* **61**: 212 (1942).
- (14) BERNARDINI, G., CACCIAPUOTI, B. N., and QUERZOLI, R.—*Ibid.* **73**: 335 (1948).
- (15) BARNÓTHY, J., and FORRÓ, M.—*Z. Phys.* **104**: 534 (1936-7).
- (16) FORBUSH, S. E.—*Phys. Rev.* **54**: 975 (1938).
- (17) FORBUSH, S. E.—*Ibid.* **70**: 771 (1946).
- (18) GAST, P. F., and LOUGHRIDGE, D. H.—*Ibid.* **58**: 583 (1940).



# HYDROGEN ATMOSPHERES IN THE ABSENCE OF THERMODYNAMIC EQUILIBRIUM

## I. THE POPULATIONS OF THE UPPER ATOMIC STATES

By R. G. GIOVANELLI\*

[*Manuscript received May 26, 1948*]

### *Summary*

From a discussion of the transition rates to and from the upper atomic states of hydrogen it is shown that for temperatures such as occur in the solar chromosphere the populations of the states  $n > 11$  are maintained largely by three-body recombination and collision ionization. Therefore, the populations of these states are distributed as in thermodynamic equilibrium with the electrons even when thermodynamic equilibrium does not exist.

These deductions are in agreement with eclipse observations.

## I. INTRODUCTION

Observations of solar and stellar spectra have been subjected, from time to time, to theoretical discussion in order to ascertain the physical properties of the emitting regions. Such discussion has been largely based on considerations implying the existence of thermodynamic equilibrium. It is known that in the solar chromosphere the temperature of the free electrons is higher than that of the radiation(1, 2) and hence it is relevant to consider the distribution of atoms among the various excited states under these conditions.

This is the first paper of a series in which such matters will be treated. An attempt is made to discuss as rigorously as possible, bearing in mind the unsatisfactory nature of our knowledge of collision cross sections, the concentration of hydrogen atoms in the highly excited states, when the distribution of electron velocities corresponds to a temperature  $\theta$ , and the atoms are exposed to diluted radiation from a black body at a temperature  $T$ .

There are, in this regard, observations which require explanation. From a discussion of the observed intensities of the Balmer lines and continuum in the spectrum of the chromosphere 670 km. above its base, Menzel and Cillié(3) have been able to derive the distribution of atoms in the highly excited states. For the lower states, their distributions depend somewhat on the adopted electron temperature, but for the states for which the principal quantum number,  $n$ , exceeds 22 they found the population to be distributed as in thermodynamic equilibrium with the electrons at the electron temperature.

\* Division of Physics, C.S.I.R.



Menzel and Cillié attempted to explain the population distribution on the basis of capture, but took into consideration only spontaneous captures and transitions. They showed that on this basis the population of the upper atomic states would have been about one-half that observed, and concluded that the excitation of the upper states is probably due to line excitation by photospheric radiation. Miyamoto(4) attempted to explain the observed population distribution by comparing the rate of collision ionization from an upper state with spontaneous transition from the upper state. He neglected, however, the effects of radiative ionization and recombination (stimulated or spontaneous), of stimulated transitions between the upper states, and of departures from a Boltzmann distribution, on the rates of transition to the upper states.

We may recall that in full statistical equilibrium, in which the principle of detailed balancing is assumed to apply, the well-known Boltzmann population distribution and the transition probabilities are such that forward and reverse transition rates between any two states are equal. In the absence of thermodynamic equilibrium, a Boltzmann distribution no longer applies, the forward and reverse rates of transition no longer balance, and the distribution of atoms among the various energy states must be found by considering the rates of all transitions to and from all the states. Such a discussion has already been given by Woolley(5) and Pannekoek(6) for transitions due solely to dilute temperature radiation. In this case hydrogen was shown to fluoresce in  $H\alpha$  and  $H\beta$ .

In the present paper, it is shown that the populations of the upper states are maintained primarily by collision ionization and three-body recombination. The only other process of significance turns out to be transition by electron collision to a neighbouring upper state. These processes thus maintain a population distribution in the upper states identical with that in thermodynamic equilibrium with the same electron concentration at the electron temperature.

## II. THE TRANSITION RATES

We first consider the rates of transition between the various states due to the absorption and emission of radiation, and to electron collisions.

### *Radiation*

(a) The number of upward transitions per unit volume per second from a lower to a higher level (denoted by  $n$  and  $r$  respectively) due to the presence of black-body radiation of temperature  $T$ , diluted by a factor  $\kappa \leq 1$ , is

$$N_n \tilde{\omega}_r \kappa A_{rn} / \tilde{\omega}_n [\exp(h\nu_{nr}/kT) - 1],$$

while the corresponding number of emissions is

$$N_r A_{rn} \{1 + \kappa / [\exp(h\nu_{nr}/kT) - 1]\},$$

where  $N_n$  and  $N_r$  represent the concentration of the atoms in the states  $n$  and  $r$ , and  $\tilde{\omega}_n$  and  $\tilde{\omega}_r$  are the respective statistical weights.

The numerical values of the coefficients  $A_{rn}$  have been computed by Slack(7) and by Menzel and Pekeris(8), and an approximate formula for  $A_{rn}$  has been given by Kramers(9).

(b) To estimate the rate of ionization by radiation, let  $\psi(\nu, n)I(\nu)d\nu dt$  be the chance that a neutral atom in state  $n$  will be ionized photoelectrically in a time  $dt$  by radiation, where  $I(\nu)d\nu$  is the quantity of radiant energy which crosses unit area normal to its path per unit solid angle per second. Then the total number of photoelectric ionizations per unit volume per second is

$$(2\pi\hbar/c^2) \sum N_n \int_{\nu_n}^{\infty} \frac{\psi(\nu, n)\nu^3 d\nu}{\exp(\hbar\nu/kT) - 1}$$

where  $\hbar\nu_n = E_n^1$  is the ionization potential.

Gaunt's value(10) for the atomic absorption coefficient  $\alpha(n, \nu)$  is

$$\alpha(n, \nu) = 2^6 \pi^4 m Z^4 \epsilon^{10} / 3 \sqrt{3} c \hbar^6 \nu^3 n^5,$$

where  $Z\epsilon$  is the nuclear charge of the atom (e.s.u.). This formula was derived for  $n > 10$ , although Gaunt hoped that the expression would give the right order of magnitude for  $n = 3$ . Menzel(11) obtained, for small quantum numbers, the formula

$$\alpha(n, \nu) = 2^9 \pi^5 m Z^4 \epsilon^{10} / 3 e^4 c \hbar^6 \nu^3 n^{4.85}.$$

The numerical factors agree to within 20 per cent., and the slight difference in the exponent of  $n$  is negligible for small quantum numbers, so that Gaunt's value may be accepted for our purposes for all values of  $n$ . Since

$$\psi(\nu, n) = 4\pi\alpha(n, \nu)/\hbar\nu,$$

the total number of ionizations of hydrogen atoms ( $Z = 1$ ) per unit volume per second is

$$\frac{2^9 \pi^5 \chi m \epsilon^{10}}{3 \sqrt{3} c^3 \hbar^6} \sum \frac{N_n}{n^5} \int_{\nu_n}^{\infty} \frac{d\nu}{\nu [\exp(\hbar\nu/kT) - 1]}.$$

For many purposes it will be sufficient to use approximations to the integral. For example, if  $\hbar\nu_n/kT > 1$ ,

$$\int_{\nu_n}^{\infty} \frac{d\nu}{\nu [\exp(\hbar\nu/kT) - 1]} \approx kT/\hbar\nu_n \exp(\hbar\nu_n/kT).$$

If  $\hbar\nu_n/kT < 1$ , then

$$\begin{aligned} \int_{\nu_n}^{\infty} \frac{d\nu}{\nu [\exp(\hbar\nu/kT) - 1]} &\approx \int_{\nu_n}^{kT/\hbar} \frac{kT d\nu}{\hbar\nu^2} + \frac{\hbar}{kT} \int_{kT/\hbar}^{\infty} \frac{d\nu}{\exp(\hbar\nu - kT)} \\ &= kT/\hbar\nu_n - 0.63 \\ &\approx kT/\hbar\nu_n. \end{aligned}$$

The rate of ionization of hydrogen atoms then becomes approximately

$$\frac{2^9 \pi^5 \chi m \epsilon^{10}}{3 \sqrt{3} c^3 \hbar^6} \sum \frac{N_n L_n}{n^5}$$

where  $L_n = kT/\hbar\nu_n \exp(\hbar\nu_n/kT)$  . . . . .  $\hbar\nu_n/kT > 1$

$L_n = kT/\hbar\nu_n$  . . . . .  $\hbar\nu_n/kT < 1$

or  $L_n \approx kT/\hbar\nu_n \exp(\hbar\nu_n/kT)$  for all  $n$ .

The summation is taken over only a finite number of quantum states, for quantized orbits are no longer possible when  $n$  becomes large. A detailed discussion of the upper limit to  $n$  has been given by a number of authors.

### Collisions

(a) Information concerning cross sections for the excitation by electron impact of hydrogen atoms from and to various energy states is very scanty. It consists of theoretical estimates of cross sections for excitation from the ground state by fast electrons, approximate cross sections for excitation to the  $2S$  and  $2P$  states by slow electrons, and experimental information on the relative cross sections for excitation by slow electrons from the ground state to the 3-, 4-, and 5-quantum states. From this we shall estimate cross sections for other transitions and it will appear that even large errors in these cross sections will not appreciably affect our conclusions.

Bethe's computations(12) of the cross sections for excitation from the ground to upper states by electrons of such energies that the first Born approximation is valid, e.g. energies greater than 1000 volts, are valuable here for extrapolation purposes.

Massey and Mohr(13) have computed the approximate cross sections of hydrogen atoms in the ground state for collisions with electrons having energies less than 45 volts, resulting in excitation to the  $2S$  and  $2P$  states. The cross sections,  $S_{1S}^{2S}$  and  $S_{1S}^{2P}$ , are roughly equal for the two transitions, and may be obtained approximately from the formula

$$S_{1S}^{2S} \approx S_{1S}^{2P} = 4.8 \times 10^{-16} / E^{\frac{1}{2}} \text{ cm.}^2,$$

where  $E$  is the energy of the incident electrons expressed in electron volts. The above formula gives a value of the total cross section,  $S_1^2$ , for excitation to the 2-quantum state which exceeds Bethe's by a factor of only about 2 for 1000-volt electrons, so that it appears reliable over a wide range of electron energies.

Ornstein and Lindeman(14) have shown experimentally that the emissions of  $H\alpha$  and  $H\beta$  are approximately constant for electron impacts at energies in the range 15-60 volts. They have further shown(15) that the ratios  $S_1^3/S_1^4$  and  $S_1^3/S_1^5$  are 3.7 and 11.5 for 30-volt electrons, as against Bethe's corresponding ratios of 2.6 and 5.4 for 1000-volt electrons, so that excitation to the higher levels seems relatively poorer for slow than for fast electrons. Bethe's value of  $S_1^2/S_1^3$  for 1000-volt electrons is 6.7. We probably make little error in assuming the ratio for slow electrons to be about 7.7, whence we obtain the following values for slow electrons

$$\begin{aligned} S_1^3 &\approx 0.130 & S_1^2 &\dots\dots\dots E > E_1^3 \\ S_1^4 &\approx 0.035 & S_1^2 &\dots\dots\dots E > E_1^4 \\ S_1^5 &\approx 0.0113 & S_1^2 &\dots\dots\dots E > E_1^5 \end{aligned}$$

where  $E_n^m$  is the difference in energy between the  $n$ - and  $m$ -quantum states.

There is no information available about the cross sections for transition by collision between excited states. We shall assume the excitation cross section for electrons whose energy is slightly above the corresponding excitation

potential,  $(S_n^r)_0$ , to be proportional to the corresponding  $B_{nr}$ , the transition probability under the influence of radiation. We shall not consider separately transitions to the various substates, as it is clear that incorrect collision transition probabilities would be obtained for forbidden transitions. For electrons having energies slightly above the excitation potential,

$$(S_n^r)_0 = \mu B_{nr} = \frac{\mu r^2}{n^2} \cdot \frac{A_{rn}}{8\pi h R^3 (n^{-2} - r^{-2})^3},$$

where  $R$  is Rydberg's number. The value of the constant,  $\mu$ , may be obtained by noting that the appropriate value of  $(S_1^2)_0$  is  $3 \cdot 5 \pi a_0^2$ . Substituting numerical values, we get

$$(S_n^r)_0 = \frac{7 \cdot 9 \times 10^{-10} \pi a_0^2 r^2 A_{rn}}{n^2 (n^{-2} - r^{-2})^3}.$$

It will be assumed that, as with  $S_1^2$ , the value of  $S_n^r$  varies as  $E^{-\frac{1}{2}}$ , where  $E$  is the energy of the colliding electron. Thus if  $E_n^r$  is the excitation potential we obtain finally

$$S_n^r = \frac{7 \cdot 9 \times 10^{-26} r^2 A_{rn}}{n^2 (n^{-2} - r^{-2})^3} \cdot \left\{ \frac{E_n^r}{E} \right\}^{\frac{1}{2}}, \dots E > E_n^r.$$

As a check on the usefulness of this formula, we find from it that for a constant  $E$ ,  $S_1^3/S_1^2$ ,  $S_1^4/S_1^2$ , and  $S_1^5/S_1^2$  are respectively 0.174, 0.062, and 0.030, of the same order of magnitude as the ratios obtained earlier.

The number of excitations from the  $n$  to the  $r$  state per unit volume per second by electrons having energies in the range  $E$  to  $E + dE$  is

$$N_n S_n^r \cdot (2E/m)^{\frac{1}{2}} N_{e,E} dE,$$

where  $N_{e,E} dE$  is the appropriate concentration of electrons, i.e.

$$N_{e,E} = 2N_e E^{\frac{1}{2}} [\exp(-E/k\theta)] / \pi^{\frac{1}{2}} (k\theta)^{3/2}.$$

Thus the rate of collision excitation is

$$[2^{3/2} N_n N_e (S_n^r)_0 (E_n^r)^{\frac{1}{2}} / (\pi m)^{\frac{1}{2}} (k\theta)^{3/2}] \int_{E_n^r}^{\infty} E^{\frac{1}{2}} \exp(-E/k\theta) \cdot dE.$$

For many purposes, we may use approximations to the integral.

Thus if  $E_n^r/k\theta > 1$

$$\int_{E_n^r}^{\infty} E^{\frac{1}{2}} \exp(-E/k\theta) \cdot dE \approx (E_n^r)^{\frac{1}{2}} k\theta \exp(-E_n^r/k\theta),$$

so that the rate of excitation is of the order of

$$2^{3/2} N_n N_e (S_n^r)_0 E_n^r \exp(-E_n^r/k\theta) / (\pi m k\theta)^{\frac{1}{2}}.$$

If, however,  $E_n^r/k\theta < 1$ ,

$$\begin{aligned} \int_{E_n^r}^{\infty} E^{\frac{1}{2}} \exp(-E/k\theta) dE &\approx \int_{E_n^r}^{k\theta} E^{\frac{1}{2}} dE + \int_{k\theta}^{\infty} (k\theta)^{\frac{1}{2}} \exp(-E/k\theta) dE \\ &\approx (k\theta)^{3/2} - \frac{2}{3} (E_n^r)^{3/2}, \end{aligned}$$

so that the rate of excitation is of the order of

$$2^{3/2} N_n N_e (S_n^r)_0 (E_n^r)^{\frac{1}{2}} (1 - \frac{2}{3} [E_n^r/k\theta]^{3/2}) / (\pi m)^{\frac{1}{2}}.$$



The number of de-excitations from the  $k$  to the  $i$  state per unit volume per second follows readily from a consideration of the principle of detailed balancing(16), according to which the rate of collision de-excitation  $N_r P_r^n$  equals the rate of collision excitation  $N_n P_n^r$  under conditions of thermodynamic equilibrium. But then  $N_n/N_r = \tilde{\omega}_n \exp(E_n^r/k\theta)/\tilde{\omega}_r = P_r^n/P_n^r$ . Thus, in the absence of thermodynamic equilibrium, but with a Maxwellian velocity distribution, the rate of de-excitation is given by

$$N_r P_r^n = N_r \tilde{\omega}_n \exp(E_n^r/k\theta) \cdot P_n^r / \tilde{\omega}_r.$$

But  $N_n P_n^r$  is the rate of excitation, derived in the previous paragraph. Thus if  $E_n^r/k\theta > 1$ ,

$$N_r P_r^n \approx 2^{3/2} N_r N_e \tilde{\omega}_n (S_n^r)_0 (E_n^r) / \tilde{\omega}_r (\pi m k \theta)^{1/2},$$

while if  $E_n^r/k\theta < 1$ , so that  $\exp(E_n^r/k\theta) \approx 1$ ,

$$N_r P_r^n \approx 2^{3/2} N_r N_e \tilde{\omega}_n (S_n^r)_0 (E_n^r) (1 - \frac{2}{3} [E_n^r/k\theta]^{3/2}) / \tilde{\omega}_r (\pi m)^{1/2}.$$

(b) The rate of ionization by collision of electrons with atoms in the  $n$ -quantum state is (17)

$$N_n \int_{E_n^i}^{\infty} S_n^i (2E/m)^{1/2} N_{e,E} dE,$$

where  $S_n^i$  is the cross section for ionization. Thus from all quantum states the rate of ionization is

$$\frac{N_e}{(\pi m)^{1/2}} \left\{ \frac{2}{k\theta} \right\}^{3/2} \sum_{n=1}^{\infty} \left\{ N_n \int_{E_n^i}^{\infty} S_n^i E \exp(-E/k\theta) \cdot dE \right\}.$$

Now  $S_n^i = \int_0^{E-E_n^i} S_n^i(\zeta, E) d\zeta$ , where  $S_n^i(\zeta, E) \approx 2\pi\epsilon^4/E(\zeta + E_n^i)^2$  (17).

Thus the rate of ionization is

$$\frac{N_e}{(\pi m)^{1/2}} \left\{ \frac{2}{k\theta} \right\}^{3/2} 2\pi\epsilon^4 \sum \left\{ N_n \int_{E_n^i}^{\infty} \left[ \int_0^{E-E_n^i} \frac{d\zeta}{E(\zeta + E_n^i)^2} \right] E \exp(-E/k\theta) \cdot dE \right\}.$$

The double integral simplifies to

$$\exp(-E_n^i/k\theta) \int_0^{\infty} \frac{\alpha \exp(-\alpha/k\theta) d\alpha}{E_n^i(E_n^i + \alpha)}.$$

The order of magnitude of the single integral is given by

$$\begin{aligned} & \int_0^k \frac{\alpha \cdot d\alpha}{E_n^i(E_n^i + \alpha)} + \int_{k\theta}^{\infty} \frac{k\theta \exp(-\alpha/k\theta) \cdot d\alpha}{E_n^i(E_n^i + k\theta)} \\ & \approx k\theta/E_n^i - \log(1 + k\theta/E_n^i) + \{k^2\theta^2/(E_n^i)^2(1 + k\theta/E_n^i)\} \exp(-1). \end{aligned}$$



If  $k\theta/E_n^i < 1$ , this is of the same order of magnitude as  $k^2\theta^2/(E_n^i)^2$ , and equals  $k^2\theta^2/(E_n^i)^2$  when  $k\theta/E_n^i \ll 1$ . If  $k\theta/E_n^i > 1$ , the order of magnitude is given by  $k\theta/E_n^i$ .

The total number of collision ionizations per unit volume per second is thus given approximately by

$$\frac{N_e}{(\pi m)^{\frac{1}{2}}} \left\{ \frac{2}{k\theta} \right\}^{3/2} 2\pi\epsilon^4 \sum N_n \cdot K_n \exp(-E_n^i/k\theta),$$

where  $K_n = (k\theta/E_n^i)^2 \dots\dots\dots k\theta/E_n^i < 1$

$K_n = k\theta/E_n^i \dots\dots\dots k\theta/E_n^i > 1$ .

It follows from the principle of detailed balancing(17) and from approximations similar to those used above, that the total number of three-body recombinations per unit volume per second is approximately

$$\frac{N_+ N_e^2}{\pi m^2} \left\{ \frac{h}{k\theta} \right\}^3 \frac{\epsilon^4}{\tilde{\omega}_i} \sum \tilde{\omega}_n K_n,$$

where  $\tilde{\omega}_i$  and  $\tilde{\omega}_n$  are the statistical weights of the ionized and the  $n$ -quantum states, respectively.

### III. POPULATIONS OF THE HIGHLY EXCITED STATES

We proceed to show that the distribution of atoms among the highly excited states ( $n$  large) is the same as in conditions of thermodynamic equilibrium at the same electron temperature and concentration.

If in thermodynamic equilibrium transitions between two states are due mainly to collisions, then a reduction of the radiation density does not affect the probability of transition between these two states. If in thermodynamic equilibrium almost all transitions from one of these states are to the other, then a reduction in the radiation density may alter the concentration of atoms in the latter state, but the ratio of the concentrations in the two states will be unchanged. We proceed to evaluate the relative probabilities for transitions from and to the upper states under the influence of collisions and radiation, and show that most of the transitions are to or from the ionized state.

(a) The ratio of the rate of collision to photoelectric ionization from the  $n$ -quantum state is

$$\frac{3\sqrt{3}N_en^5c^3h^6 \exp(-E_n^i/k\theta)}{2^6\pi^4(2\pi)^{\frac{1}{2}}\chi(mk\theta)^{3/2}\epsilon^6L_n} \cdot K_n.$$

Considering only upper states, so that  $E_n^i/k\theta$  and  $E_n^i/kT$  are both less than unity, the above expression becomes

$$\frac{3\sqrt{3}N_en^5c^3h^6 \exp(-E_n^i/k\theta)}{2^6\pi^4(2\pi)^{\frac{1}{2}}\chi(mk)^{3/2}\theta^{\frac{1}{2}}T\epsilon^6}$$

which, on substituting numerical values and replacing the exponential by unity, reduces to  $1.4 \times 10^{-9} N_e n^5 / \chi \theta^{\frac{1}{2}} T$ . If we put  $\theta = \chi T = 5 \times 10^3$ , this becomes  $3.9 \times 10^{-15} N_e n^5$ . In the chromosphere,  $N_e$  is of the order of  $10^{10}$ , so the ratio is of the order of  $4 \times 10^{-5} n^5$ , and considerably exceeds unity when  $n$  exceeds 10. It is readily seen that the above expression applies for these values of  $n$ , since

$E_n^i/k\theta = hcR/k\theta n^2 = 1.58 \times 10^5/n^2\theta$  and, for  $\theta = 5 \times 10^3$ , is less than unity when  $n \geq 7$ . Collision ionizations are thus much more frequent than photoelectric ionizations from the upper states.

(b) The ratio of collision ionizations from the  $n$ -quantum state to transitions from this state to other discrete states with the emission or absorption of radiation is

$$\frac{N_e}{(\pi m)^{\frac{1}{2}}} \left\{ \frac{2}{k\theta} \right\}^{3/2} 2\pi\epsilon^4 K_n \exp(-E_n^i/k\theta) \cdot \frac{\sum_{l < n} A_{nl} \{1 + \kappa/[\exp(E_n^l/kT) - 1]\} + \sum_{l > n} (l/n)^2 A_{ln} \kappa/[\exp(E_n^l/kT) - 1]}{}$$

If  $n$  is large, so that  $E_n^i/k\theta < 1$ , then  $K_n = k\theta/E_n^i$ . The value of the above ratio may be estimated by introducing Kramers' approximation to  $A_{nl}$  (9), viz.

$$A_{nl} = \frac{2^8 \pi \epsilon^2 R^2}{3 \sqrt{3} m c l^3 n^5 (l^{-2} - n^{-2})} \\ = 1.58 \times 10^{10} / l^3 n^5 (l^{-2} - n^{-2}),$$

and noting that the denominator approximates to

$$\sum_{l < n} A_{nl} + \sum_{m < l < n} A_{nl} \kappa kT/E_n^l + \sum_{l > n} (l/n)^2 A_{ln} \kappa kT/E_n^l,$$

where  $E_n^m/kT = 1$ . If  $n$  is large, we may put  $\exp(-E_n^i/k\theta)$  equal to unity and, inserting numerical constants, the ratio then reduces to

$$\frac{1.52 \times 10^{-15} N_e}{\theta^{\frac{1}{2}} [3.45 \times 10^{-1} n^{-7} \sum_{l < m} l^{-3} (l^{-2} - n^{-2})^{-1} + 2.18 \times 10^{-6} \kappa T n^{-3} \sum_{m < l < n} l (n^2 - l^2)^{-2} \\ + 2.18 \times 10^{-6} \kappa T n^{-3} \sum_{l > n} l (l^2 - n^2)^{-2}]}$$

Of the terms in the denominator, the first is of the order of  $3 \times 10^{-1} n^{-7}$ . The

summation in the second term is approximately  $\int_m^{n-1} l (l^2 - n^2)^{-2} dl$ , which in

turn approximates to  $1/4n$  for  $n$  large; the second term is thus of the order of  $5 \times 10^{-7} \kappa T n^{-4}$  for  $n$  large. The third term approximates to the same value, so that the denominator is of the order of

$$\theta^{\frac{1}{2}} (3 \times 10^{-1} n^{-7} + 2 \times 5 \times 10^{-7} \kappa T n^{-4}).$$

If  $\kappa T$  is  $5 \times 10^3$  °K., this expression is of the order of

$$\theta^{\frac{1}{2}} (3 \times 10^{-1} n^{-7} + 5 \times 10^{-3} n^{-4}).$$

Of these, the second term is the larger when  $n$  exceeds 5; so that it alone needs to be considered in the range of  $n$  in which we are interested. The ratio is then of the order of

$$\frac{1.5 \times 10^{-15} N_e}{10^{-6} \kappa T \theta^{\frac{1}{2}} n^{-4}}.$$

Putting  $\kappa T = \theta = 5 \times 10^3$ ,  $N_e = 10^{10}$ , this becomes  $4 \times 10^{-5} n^4$ , and exceeds unity when  $n \geq 11$ .

Owing to the approximate nature of the computations, not too much reliance should be placed on the actual numerical value obtained here for  $n$ ; it is safe to conclude, however, that for  $n$  not much larger than 11, collision ionizations occur much more rapidly than transitions by the emission or absorption of radiation.

(c) The ratio of the rate of ionization by collision from a state  $n$  to the rate of transition to other discrete states from the same state by collision is

$$\frac{2\pi\epsilon^4 \exp(-E_n^i/k\theta) \cdot (E_n^i/k\theta)^{\frac{1}{2}}}{\sum_{m \leq l} \frac{m^2 (S_m^n)_0 E_m^n}{n^2 (k\theta)^{\frac{1}{2}}} + \sum_{l < m < n} \frac{m^2 (S_m^n)_0 (E_m^n)^{\frac{1}{2}} (1 - \frac{2}{3} [E_m^n/k\theta]^{3/2})}{n^2} + \sum_{m > n} \frac{(S_n^m)_0 (E_n^m)^{\frac{1}{2}} (1 - \frac{2}{3} [E_n^m/k\theta]^{3/2})}{n^2}}$$

where  $E_n^i/k\theta < 1$ ,  $E_l^n/k\theta \approx 1$ . But using Kramers' value for  $A_{mn}$ ,

$$(S_n^m)_0 = 1 \cdot 1 \times 10^{-15} / n^5 m^3 (n^{-2} - m^{-2})^4;$$

also,  $E_n^m = hcR(n^{-2} - m^{-2}) = 2 \cdot 19 \times 10^{-11} (n^{-2} - m^{-2})$ .

Introducing these values, we find, after some reduction, that the ratio becomes

$$\frac{1 \cdot 18 \times 10^{-3} n^7 \exp(-E_n^i/k\theta) \cdot (\theta)^{\frac{1}{2}}}{\sum_{m \leq l} 1 \cdot 87 \times 10^{-3} / m^3 (m^{-2} - n^{-2})^3 \theta^{\frac{1}{2}} + \sum_{l < m < n} 4 \cdot 68 \times 10^{-6} (1 - \frac{2}{3} [E_m^n/k\theta]^{3/2}) / m^3 (m^{-2} - n^{-2})^{7/2} + \sum_{m < n} 4 \cdot 68 \times 10^{-6} (1 - \frac{2}{3} [E_n^m/k\theta]^{3/2}) / m^3 (n^{-2} - m^{-2})^{7/2}}$$

Of the terms in the denominator, the first is of the order of  $2 \times 10^{-3} \sum m^3/\theta$ , and as  $l$  is of the order of 4 if  $\theta$  is of the order of  $10^4$  °K., this term is about

$$2 \times 10^{-3}. \text{ The second term approximates to } 5 \times 10^{-6} \int_l^{n-1} dm / m^3 (m^{-2} - n^{-2})^{7/2},$$

if we put  $1 - \frac{2}{3} [E_m^n/k\theta]^{3/2} \approx 1$ . Its value is approximately  $5 \times 10^{-6} n^{15/2} / 5 \times 2^{5/2}$ , about  $10^{-7} n^{15/2}$ . The third term has a similar value. For values of  $n$  for which the above expressions apply, the first term is negligibly small, and collision transitions are mainly to the upper states, particularly the states near  $n$ .

The ratio of the rate of ionization to the rate of transition by collision is thus of the order of  $5 \times 10^3 / n^3 \theta^{\frac{1}{2}}$ ; for  $\theta = 10^4$ , the rate of ionization exceeds the rate of transition by collision for all *quantized orbits* for which the above analysis applies, i.e. for  $n > 4$ . In any case, as transitions by collision are mainly to neighbouring states, they will not disturb conditions appropriate to thermodynamic equilibrium at the same electron temperature and concentration.

(d) We have seen that transitions from upper states are mainly ionizations by collision. It will now be shown that transitions to upper states are mainly three-body recombinations. The proof is not as straightforward as in the discussion on transitions from state  $n$ , as we must first make some estimate of the upper limit to the various values of  $N_n$ .

If the density of radiation be decreased below that in thermodynamic equilibrium at the temperature  $\theta$  of the electrons, the ratio of the probability of an upward transition to that of a downward transition is decreased, and for a given number of atoms in the ground state the population of the upper states is reduced. The electron concentration is similarly lowered. If, however, as in the sun's atmosphere, the electron concentration can be deduced from observation, a lower radiation density implies a higher concentration of neutral atoms, so that the importance of transitions between the discrete states may be greater, relative to three-body recombinations. We therefore commence by giving an

approximate ionization formula for conditions which depart from those of thermodynamic equilibrium. A fuller discussion of the ionization will be given in a subsequent paper.

We assume here that the electron concentration depends on ionizations from, and recombinations to, the ground state. The recombinations are mainly spontaneous. Equating rates of ionization and recombination\* we find

$$4.4 \times 10^{-10} N_e N_1 \theta^{\frac{1}{2}} \exp(-1.58 \times 10^5/\theta) + 5.0 \times 10^4 \kappa N_1 T \exp(-1.58 \times 10^5/T) \\ = 2.07 \times 10^{-11} N_e^2 \theta^{\frac{1}{2}},$$

where the first and second terms on the left-hand side give the rates of ionization by collision and radiation respectively. The first or second term is the more important according as

$$8.8 \times 10^{-15} N_e \theta^{\frac{1}{2}} \exp(-1.58 \times 10^5/\theta) \gtrless \kappa T \exp(-1.58 \times 10^5/T).$$

For  $N_e = 10^{10}$ ,  $\kappa = 1$ ,  $\theta = 10^4$ , the condition is roughly  $T \leq 5400$  °K. For the lower radiation densities the equation of ionization is roughly

$$N_e/N_1 \approx 21\theta \exp(-1.58 \times 10^5/\theta),$$

while for the higher radiation densities it is

$$N_e^2/N_1 \approx 2.4 \times 10^{15} \kappa T \theta^{\frac{1}{2}} \exp(-1.58 \times 10^5/T).$$

In thermodynamic equilibrium at a temperature  $\theta$ , the concentrations of electrons and atoms in the  $j$  state are given by combining Saha's and Boltzmann's relations :

$$N_j = \frac{\tilde{\omega}_j N_e^2 h^3 \exp(E_j^i/k\theta)}{\tilde{\omega}_1 (2\pi m k \theta)^{3/2}} \\ = 4.15 \times 10^{-16} N_e^2 \tilde{\omega}_j \{\exp(E_j^i/k\theta)\} / \tilde{\omega}_1 \theta^{3/2}.$$

For a given electron concentration, the effect of decreasing the intensity of radiation is to require a population in the ground state higher by a factor of  $M_1$  than that required under conditions of thermodynamic equilibrium. For low radiation densities

$$M_1 \approx 1.15 \times 10^{14} \theta^{\frac{1}{2}} / N_e$$

or about  $10^6$  if  $\theta = 10^4$ ,  $N_e = 10^{10}$ ; for the higher radiation densities,  $M_1$  is less, and is given by

$$M_1 \approx (\theta/\kappa T) \exp(1.58 \times 10^5[1/T - 1/\theta]).$$

The populations of the excited states with respect to that of the ground state will be much less than in thermodynamic equilibrium at  $\theta$ . Suppose, for example, that the population of state  $n$  is maintained by excitation (radiation or collision) direct from the ground state, de-excitation being by spontaneous transition direct to the ground state. (This simplification does not necessarily give a very good approximation for the population of the upper states, but for the lower states it is sufficient for present purposes.) Then, equating rates of excitation and de-excitation, we find, when  $E_1^n/k\theta > 1$ ,

$$6.9 \times 10^{-15} N_1 N_e A_{n1} \exp(-E_1^n/k\theta) / \theta^{\frac{1}{2}} + N_1 n^2 \kappa A_{n1} \exp(-E_1^n/kT) = N_n A_{n1}.$$

\* The rate of recombination is a well-known expression; it is derived explicitly in the following paper.



The collision excitations are given by the first term, and exceed radiative excitations when

$$6 \cdot 9 \times 10^{-15} N_e \exp(-E_1^n/k\theta) / \theta^{\frac{1}{2}} > n^2 \kappa \exp(-E_1^n/kT).$$

For  $n=2$ ,  $\kappa=1$ ,  $N_e=10^{10}$ ,  $\theta=10^4$ , we find this condition to be  $T < 4500^\circ \text{K}$ . For low radiation densities the concentration in the excited state is roughly

$$N_n = 6 \cdot 9 \times 10^{-15} N_1 N_e \exp(-E_1^n/k\theta) / \theta^{\frac{1}{2}},$$

while for higher radiation densities it is

$$N_n = N_1 n^2 \kappa \exp(-E_1^n/kT).$$

The ratio of the population to that for an electron concentration  $N_e$  in thermodynamic equilibrium at temperature  $\theta$  is  $M_n$ , where for low radiation densities

$$M_n \approx 0 \cdot 4 / n^2,$$

and for the higher radiation densities

$$M_n \approx (\theta/T) \exp(E_n^i/k) \left( \frac{1}{\theta} - \frac{1}{T} \right).$$

Thus, except in the ground state, the concentrations remain of the same magnitude, for constant  $N_e$ , when the radiation intensities are reduced. This argument, it should be pointed out, does not necessarily apply to the metastable  $2S$  state, and refinements including the effects of other transitions will be discussed in a following paper.

(e) We are now in a position to examine rates of transition to the upper states. Since photoelectric ionizations from upper states occur much less rapidly than collision ionizations in thermodynamic equilibrium, the reverse process of photoelectric recombination may also be neglected. This is particularly so if the radiation density be reduced.

(f) The ratio of the rate of three-body recombination to an upper state  $n$  to the rate of transition to that state by the absorption or emission of radiation is

$$\frac{2n^2 N_e^3 h^3 \epsilon^4 / \pi (mk\theta)^2 E_n^i}{\sum_{l < n} \frac{N_l \cdot 2n^2 \kappa A_{nl}}{2l^2 [\exp(E_l^n/kT) - 1]} + \sum_{l > n} N_l A_{ln} \left\{ 1 + \frac{\kappa}{\exp(E_n^l/kT) - 1} \right\}}$$

if  $N_e = N_+$ . But

$$N_l = l^2 N_e^2 M_l h^3 \exp(E_l^i/k\theta) / (2\pi mk\theta)^{3/2}.$$

Introducing these values for  $N_l$ , and inserting Kramers' values for  $A_{nl}$ , the above ratio reduces to

$$\frac{3\sqrt{3} N_e \epsilon^2 m^{\frac{1}{2}} c n^5 / 2^{11/2} (\pi k\theta)^{\frac{1}{2}} E_n^i R^2}{\sum_{l < n} \frac{\kappa M_l \exp(E_l^i/k\theta)}{l(l^2 - n^2) L_n} + \sum_{l > n} \frac{M_l \exp(E_l^i/k\theta)}{l^3 (l^2 - n^2)} \left\{ 1 + \frac{\kappa kT}{E_n^i} \right\}}$$

where  $L_n \approx \exp(E_l^n/kT)$ , .....  $E_l^n/kT > 1$

$L_n \approx E_l^n/kT$  .....  $E_l^n/kT < 1$ .



Inserting numerical values, the ratio becomes

$$\frac{4 \cdot 4 \times 10^{-15} N_e n^7 / \theta^{\frac{1}{2}}}{\sum_{l \leq r} \frac{\kappa M_l \exp(E_l^i / k\theta - E_l^n / kT)}{l(l^{-2} - n^{-2})} + \sum_{r < l < n} \frac{\kappa T M_l \exp(E_l^i / k\theta)}{1 \cdot 58 \times 10^5 l(l^{-2} - n^{-2})^2} + \sum_{l > n} \frac{M_l \exp(E_l^i / k\theta)}{l^3(l^{-2} - n^{-2})} (1 + \kappa kT / E_n^l)}$$

where  $E_1^r / kT \approx 1$ .

It follows from (b) of this section and from the principle of detailed balancing that the rate of three-body recombination to state  $n \geq 11$  exceeds the rate of transition to this state, by radiative processes, from all states for which  $M_l$  is unity. We shall assume the result that  $M_l$  is unity for all states above  $m$ , where  $m$  is a small number, and so need concern ourselves solely with transitions from states 1 to  $m$ . The first term in the denominator is of the order of  $\kappa M_1 \exp(E_1^i / k\theta - E_1^n / kT)$ . For the higher radiation densities, such that ionizations from the ground state are due mainly to the absorption of radiation, this term has a value of  $(\theta/T) \exp(E_n^i / kT)$ . For constant  $\theta$ , this expression increases for decreasing  $T$ . When  $T$  becomes so low that ionizations from the ground state are predominantly due to electron collisions, the corresponding term decreases with  $T$ . The approximate maximum value is obtained by noting that the exponential is of the order of unity if  $n$  is large, and  $\theta/T$  is also of the order of unity when collision and radiative ionizations from the ground states occur at equal rates. The ratio of the numerator to this term of the denominator is thus of the order of  $10^{-15} N_e n^7 / \theta^{\frac{1}{2}}$ ; putting  $N_e = 10^{10}$ ,  $\theta = 10^4$  °K., this ratio is  $10^{-7} n^7$  and exceeds unity when  $n$  exceeds 10.

Since  $M_l (l > 1)$  is of the order of unity, then radiative transitions from states above  $r$  decrease as  $T$  decreases; so that the ratio of the rate of three-body recombination to state  $n$  to the rate of radiative transition from states above  $r$  will exceed unity for conditions not very different from those found in (b), i.e. approximately  $n > 11$ .

(g) In estimating the relative importances of three-body recombinations and transitions to state  $n$  by collision, we may neglect transitions from the upper states, as in (f), and consider only collision excitations from states for which  $E_l^n / k\theta > 1$ . The ratio of the rate of three-body recombination to the rate of such collision excitation is

$$\frac{2n^2 N_e^3 h^3 \varepsilon^4 / \pi (mk\theta)^2 E_n^i}{\sum 2^{3/2} N_l N_e (S_l^n)_0 E_l^n \exp(-E_l^n / k\theta) / (\pi mk\theta)^{\frac{1}{2}}},$$

which on substituting numerical values reduces to

$$\frac{0 \cdot 64 n^7 \exp(-E_n^i / k\theta)}{\sum M_l / l^3 (l^{-2} - n^{-2})^3}.$$

The value of the denominator is closely equal to the value of the term for  $l=1$ ; the ratio for the higher radiation densities is thus approximately

$$0 \cdot 64 (\kappa T / \theta) n^7 \exp(E_1^n / k\theta - E_1^i / kT).$$

This expression holds only when the radiation density is such as to produce a rate of ionization by radiation exceeding that by collision. For lower radiation

densities, the ratio of the rate of three-body recombination to collision excitation from the ground state has a fairly constant minimum value of

$$5.6 \times 10^{-15} n^7 \exp(-E_n^i/k\theta) \cdot N_e/\theta^{\frac{1}{2}}.$$

For example, if  $\theta=10^4$ ,  $N_e=10^{10}$ , the ratio acquires a value of  $5.6 \times 10^{-7} n^7$ , and exceeds unity when  $n > 7$ .

(h) We have now shown that, for conditions such as are likely to occur in the sun's atmosphere, collision ionizations and three-body recombinations to an upper state  $n$  are more important than transitions (to and from this upper state) which involve the lower levels; that they are more important than transitions to neighbouring levels (especially those transitions which involve the emission or absorption of radiation); and, on the assumption that the populations of the upper states equal those in conditions of thermodynamic equilibrium corresponding to the same electron concentration and temperature, that they are more important than transitions from the neighbouring levels. It follows that the populations of the upper levels are in fact distributed as if the atoms were in thermodynamic equilibrium with the electrons.

Our investigations have suggested that other processes, namely, transitions to neighbouring upper states under the stimulation of radiation, and radiative excitations from the ground state, may be equally as important as collision ionization and three-body recombination with  $n \lesssim 11$ ; so that the populations may be expected to depart from the thermodynamic equilibrium values when  $n$  is of this order. These departures increase as  $n$  decreases. The various numerical values found for  $n$  for which our conclusions apply have all been obtained for specific values of  $\theta$ ,  $T$ ,  $\kappa$ , and  $N_e$ , and will vary with these quantities. On the other hand, the relative importance of the collision ionization and the three-body recombination processes increases as a high power of  $n$ , so that over a wide range of conditions rather similar results will be obtained.

(i) From the above treatment of a hydrogen atmosphere, some inferences of a similar nature may be made for atmospheres of any other gas or of a mixture of gases. The case of a mixture of gases is important in solar and stellar atmospheres. Let the concentration of electrons be  $N_e$ , the concentration of ions of an element  $A$  be  $N_{A+}$ . Then the rate of three-body recombination is proportional to  $N_e^2 N_{A+}$ , while the rate of collision ionization from a given upper state  $n$  is proportional to  $N_e N_n$ . If the population of the upper state depends on the balancing of the rates of collision ionization and three-body recombination, then  $N_n$  is proportional to  $N_e N_{A+}$ . As some of the electrons may be due to the ionization of other elements,  $N_e$  may exceed  $N_{A+}$ . Thus for a given concentration of ions of one type, the populations of the upper states of the corresponding neutral atom will exceed those in a pure gas in thermodynamic equilibrium at the temperature of the electrons and with the same ion concentration, in the ratio  $N_e/N_{A+}$ . In other words, Saha's ionization formula applies to the upper states when the temperature used is that of the electrons.

## IV. APPLICATIONS TO THE SOLAR CHROMOSPHERE

The above conclusions may now be compared with Menzel and Cillié's observed populations. We have found, in general, that collision excitations and three-body recombinations are the most important processes in maintaining the populations of states for which  $n > 11$ , so that for  $n \gg 11$  the distribution should be the same as in thermodynamic equilibrium with the electrons. From observation, this distribution in fact applies when  $n \geq 22$ . With an assumed electron temperature of 10,000 °K., the observed populations of the states  $n=17, 11$ , and 5 were found to be about 75, 60, and 30 per cent. of the thermodynamic equilibrium values, although, as Menzel and Cillié pointed out, their discussions neglected the effects of self-absorption which would have raised the computed populations somewhat for the lower states. These observations are in good agreement with the predictions of the present paper.

In a subsequent paper, we shall discuss the problem of the populations of the lower and ground states.

## V. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Physics, C.S.I.R.

## VI. REFERENCES

- (1) CILLIÉ, G. G., and MENZEL, D. H.—*Harv. Obs. Circ.* No. 410 (1935).
- (2) REDMAN, R. O.—*Mon. Not. Roy. Astr. Soc.* **102**: 140 (1942).
- (3) MENZEL, D. H., and CILLIÉ, G. G.—*Astrophys. J.* **85**: 88 (1937).
- (4) MIYAMOTO, S.—*Mem. Coll. Sci. Kyoto* **22A**: 1 (1939).
- (5) WOOLLEY, R. v. d. R.—*Mon. Not. Roy. Astr. Soc.* **94**: 631 (1934).
- (6) PANNEKOEK, A.—*Ibid.* **95**: 725 (1935).
- (7) SLACK, F. G.—*Phys. Rev.* **31**: 527 (1928).
- (8) MENZEL, D. H., and PEKERIS, C. L.—*Mon. Not. Roy. Astr. Soc.* **96**: 77 (1935).
- (9) KRAMERS, H. A.—*Phil. Mag.* **44**: 836 (1923).
- (10) GAUNT, J. A.—*Philos. Trans. A* **229**: 163 (1930).
- (11) MENZEL, D. H.—*Lick Obs. Publ.* **17**: 1 (1931).
- (12) BETHE, H., quoted in MOTT, N. F., and MASSEY, H. S. W.—“The Theory of Atomic Collisions”, p. 183. (Clarendon Press, 1933.)
- (13) MASSEY, H. S. W., and MOHR, C. B. O.—*Proc. Roy. Soc. A* **132**: 605 (1931).
- (14) ORNSTEIN, L. S., and LINDEMAN, H.—*Proc. Acad. Sci. Amst.* **33**: 1097 (1930).
- (15) ORNSTEIN, L. S., and LINDEMAN, H.—*Z. Phys.* **67**: 1 (1931).
- (16) FOWLER, R. H.—“Statistical Mechanics”, p. 431. (Cambridge University Press, 1929.)
- (17) FOWLER, R. H.—*Ibid.* p. 438.

# HYDROGEN ATMOSPHERES IN THE ABSENCE OF THERMODYNAMIC EQUILIBRIUM

## II. THE POPULATIONS OF THE LOWER ATOMIC STATES

By R. G. GIOVANELLI\*

[*Manuscript received June 9, 1948*]

### *Summary*

The populations of the hydrogen substates of principal quantum number 1, 2, or 3 are calculated for a range of electron concentrations from  $10^8$  to  $10^{13}$  per cc. and electron temperatures from  $7.5 \times 10^3$  to  $1.0 \times 10^5$  °K., in the presence of radiation from a black body at a temperature of  $5 \times 10^3$  °K., diluted by a factor of 1, 0.5, or 0.

## I. INTRODUCTION

In a previous paper(1) the populations of the upper states of hydrogen were examined under a range of conditions which depart from those of thermodynamic equilibrium. These conditions were defined by the presence of dilute radiation from a black body at a temperature  $T$ , while the distribution of electron velocities was assumed to correspond to a temperature  $\theta$  greater than  $T$ . The populations of the upper atomic states of hydrogen were shown to depend only on the electron temperature and electron concentration, and to be as in thermodynamic equilibrium with the electrons. We proceed here to discuss the populations of the states of which the principal quantum number  $n$  is 1, 2, or 3 when the electron temperature and concentration and the spectral distribution and the density of radiant energy are given. Results of calculations are tabulated for electron concentrations from  $10^8$  to  $10^{13}$  per cc., electron temperatures from  $7.5 \times 10^3$  to  $1.0 \times 10^5$  °K., and black-body radiation, temperature  $5 \times 10^3$  °K., with a dilution factor  $\kappa=1, 0.5$ , or 0.

As the populations of the various states cannot be observed directly, but only indirectly by way of the radiation emitted or absorbed during transfers between the states, a theoretical study of the intensity distribution in the spectrum is essential before we can discuss observations of stellar atmospheres in terms of the results presented here. This forms the subject of a following paper.

The methods used are necessarily approximate, because of the difficulties in solving a large group of simultaneous equations, in each of which the rates at which atoms arrive and depart from a given state are equated. Moreover, the transition probabilities are in many cases known only approximately or by inference. The analysis, nevertheless, yields results which should be of the correct order of magnitude for the populations of the three lowest states and for the concentration of the ions.

\* Division of Physics, C.S.I.R.



## II. THE TRANSITION RATES

*Radiation*

The number of transitions per unit volume per second from a lower level,  $j$ , to a higher level,  $l$ , due to the presence of black-body radiation of temperature  $T$  diluted by a factor  $\kappa \leq 1$ , is

$$N_j \tilde{\omega}_j \kappa A_{l,j} / \tilde{\omega}_j [\exp(E_j/kT) - 1].$$

The corresponding number of emissions is

$$N_l A_{l,j} \{1 + \kappa / [\exp(E_j/kT) - 1]\},$$

$N_j$  and  $N_l$  being the concentrations of atoms in the states  $j$  and  $l$  respectively, and  $\tilde{\omega}_j$  and  $\tilde{\omega}_l$  the statistical weights. The values of the coefficients  $A_{l,j}$  are well known(2).

The rate of ionization from a state of principal quantum number  $n$  is given by

$$\frac{2^9 \pi^5 \chi m \epsilon^{10} N_n}{3 \sqrt{3} c^3 h^6 n^5} \int_{E_n^i}^{\infty} \frac{dE}{E [\exp(E/kT) - 1]}.$$

In most cases of interest to us,  $E_n^i/kT \gg 1$ , so that the integral is approximately  $-Ei(-E_n^i/kT)$ , where

$$-Ei(-x) = \int_x^{\infty} \frac{\exp(-t)}{t} dt.$$

Thus, inserting numerical values, the rate of ionization is

$$7.85 \times 10^9 \kappa (N_n/n^5) \{-Ei(-E_n^i/kT)\}.$$

No distinction is made, in this formula, between the atoms which are initially in the various substates; when considering ionization from a substate, the appropriate value of  $N_j$  will be used.

We may note that when  $E_n^i/kT > 1$ , the exponential integral is given, to within a factor of 2, by  $kT/E_n^i \exp(E_n^i/kT)$ . Since  $E_n^i = 2.19 \times 10^{-11}/n^2$ , the rate of ionization is approximately

$$5.0 \times 10^4 \kappa N_n T / n^3 \exp(E_n^i/kT).$$

The rate of recombination may be derived from expressions given by Fowler(3). After inserting the energy distribution functions for the radiation and for the electrons into these expressions, it is found that the rate of recombination to a state of principal quantum number  $n$  is

$$\frac{2^7 \pi^4}{3} \left(\frac{2}{3\pi}\right)^{\frac{1}{2}} \frac{N_+ N_e \epsilon^{10}}{(k\theta)^{3/2} c^3 h^3 m^{\frac{1}{2}} n^3} \int_0^{\infty} \frac{\exp(-E/k\theta)}{(E - E_n^i)} \left(1 + \frac{\kappa}{\exp\{(E + E_n^i)/kT\} - 1}\right) dE.$$

Induced captures may be neglected if  $\kappa / (\exp\{E_n^i/kT\} - 1) \ll 1$ , as will usually be the case. Inserting numerical values, the rate of recombination is found to be

$$3.28 \times 10^{-6} (N_+ N_e / n^3 \theta^{3/2}) \{-Ei(-E_n^i/k\theta)\} \exp(E_n^i/k\theta).$$

If  $E_n^i/k\theta > 1$ , this is given, to within a factor of 2, by  $2.07 \times 10^{-11} N_+ N_e / n \theta^{\frac{1}{2}}$ .



Since these recombinations will presumably be distributed among the substates according to their statistical weights, and as the weight of the state is  $2n^2$ , the rate of spontaneous recombination to a substate  $j$  is obtained by multiplying the above expressions by  $\tilde{\omega}_j/2n^2$ .

### Collisions

Cross sections for the excitation of various states of hydrogen by collision with slow electrons are not accurately known. From available information, we can do little better than accept the excitation rate used in (1), viz. the rate of excitation from state  $n$  to state  $r$  is approximately

$$2^{3/2}N_nN_e(S_n^r)_0E_n^r \exp(-E_n^r/k\theta)/(\pi mk\theta)^{1/2}, \dots\dots\dots E_n^r/k\theta > 1,$$

or

$$2^{3/2}N_nN_e(S_n^r)_0(E_n^r)^{1/2}(1 - \frac{2}{3}[E_n^r/k\theta]^{3/2})/(\pi m)^{1/2}, \dots\dots\dots E_n^r/k\theta < 1,$$

where

$$(S_n^r)_0 \approx \frac{7 \cdot 9 \times 10^{-10} \pi a_0^2 r^2 A_{r,n}}{n^2(n^2 - r^2)^3},$$

and  $a_0$  is the radius of the first Bohr orbit. These expressions become, on substituting numerical data,

$$\frac{6 \cdot 9 \times 10^{-15} N_n N_e r^2 A_{r,n} \exp(-E_n^r/k\theta)}{n^2(n^2 - r^2)^2 \theta^{1/2}}, \dots\dots\dots E_n^r/k\theta > 1$$

and

$$\frac{1 \cdot 72 \times 10^{-17} N_n N_e r^2 A_{r,n} [1 - 4 \cdot 2 \times 10^7 (n^2 - r^2)^{3/2} / \theta^{3/2}]}{n^2(n^2 - r^2)^{5/2}}, \dots\dots\dots E_n^r/k\theta < 1.$$

These expressions apply to the rates of transfer between states. In the absence of better information, we assume, as has been computed(4) for  $1S \rightarrow 2S$  and  $1S \rightarrow 2P$ , that the rates of transfer to the various substates are approximately equal. Thus the rate of collision excitation to each of the  $r$  substates of state  $r$  is  $1/r$  times the above expressions.

Superelastic collisions occur between states  $l$  and  $j$  at a rate

$$\frac{2^{3/2}N_lN_e\tilde{\omega}_j(S_j^l)_0E_j^l}{\tilde{\omega}_l(\pi mk\theta)^{1/2}}, \dots\dots\dots E_j^l/k\theta > 1$$

or

$$\frac{2^{3/2}N_lN_e\tilde{\omega}_j(S_j^l)_0(E_j^l)^{1/2}(1 - \frac{2}{3}[E_j^l/k\theta]^{3/2})}{\tilde{\omega}_l(\pi m)^{1/2}}, \dots\dots\dots E_j^l/k\theta < 1.$$

Such collisions are of importance for the transition  $2S \rightarrow 1S$ , for which the rate is  $1 \cdot 16 \times 10^{-5} N_{2S} N_e / \theta^{1/2}$ ,  $E_1^2/k\theta > 1$ , and may be of some small importance for the transitions  $3 \rightarrow 2$ , for which the rates are

$$8 \cdot 0 \times 10^{-13} N_{3x} N_{3y} \tilde{\omega}_{3x} A_{3,2} / \tilde{\omega}_{3x} \theta^{1/2}, \dots\dots\dots E_2^3/k\theta > 1.$$

The rate of ionization by collision from state  $n$  is given approximately by

$$\frac{N_e}{(\pi m)^{\frac{1}{2}}} \left( \frac{2}{k\theta} \right)^{3/2} 2\pi e^4 N_n K_n \exp(-E_n^i/k\theta),$$

where  $K_n \approx (k\theta/E_n^i)^2, \dots, E_n^i/k\theta > 1$

$$K_n \approx k\theta/E_n^i, \dots, E_n^i/k\theta < 1.$$

Substituting numerical values, the rate of collision ionization becomes

$$4.4 \times 10^{-10} N_e N_n n^4 \theta^{\frac{1}{2}} \exp(-E_n^i/k\theta), \dots, E_n^i/k\theta > 1$$

and

$$6.9 \times 10^{-5} N_e N_n n^2 \theta^{-\frac{1}{2}} \exp(-E_n^i/k\theta), \dots, E_n^i/k\theta < 1.$$

The rate of three-body recombination to a state  $j$  is

$$\frac{N_+ N_e^2}{\pi m^2} \left( \frac{h}{k\theta} \right)^3 \frac{\varepsilon^4 \tilde{\omega}_j K_j}{\tilde{\omega}_i},$$

where  $K_j \approx (k\theta/E_j^i)^2, \dots, E_j^i/k\theta > 1$

or  $K_j \approx k\theta/E_j^i, \dots, E_j^i/k\theta < 1.$

This reduces to

$$\frac{8.8 \times 10^{-26} N_+ N_e^2 \tilde{\omega}_j n^4}{\theta \tilde{\omega}_i}, \dots, E_j^i/k\theta > 1$$

or

$$\frac{1.38 \times 10^{-20} N_+ N_e^2 \tilde{\omega}_j n^2}{\theta^2 \tilde{\omega}_i}, \dots, E_j^i/k\theta < 1.$$

Three-body recombinations to states  $n=1$  or  $2$  are found to be of little importance compared with spontaneous recombinations and will be neglected, but they may be of importance for recombinations to  $n=3$ .

There is a very important type of collision, about the probability of which there appears to be no information, viz. an elastic collision in which the atom is converted from a substate  $j$  of an energy level to another substate  $l$  of the same level, e.g.  $2S \rightarrow 2P$ . Let the cross section for electrons of velocity  $u$  be  $X_j^l$ . Then the total number of such collisions, per unit volume per second, is

$$N_j \int u X_j^l dN_e.$$

Suppose, for example, that  $u X_j^l$  is independent of  $u$ , and is  $10^{-8}$  cm.<sup>3</sup> per second. Then the rate of such collisions is  $10^{-8} N_j N_e$  per unit volume per second. It seems likely that elastic collisions of this type will be no more important than super-elastic or inelastic collisions in removing atoms from a particular substate, and we neglect them in the following analysis. However, if the cross section is very much larger than  $10^{-8}/u$  cm.<sup>2</sup>, the results tabulated below may contain errors which will be most significant for high electron concentrations. These errors would be most serious in the case of the metastable  $2S$  state, whose population would thereby be reduced.

## III. THE EQUILIBRIUM EQUATIONS

We shall restrict this discussion to populations maintained by transitions between the 1, 2, 3, and ionized states alone; transitions between any one of these states and the higher states are thus assumed to be negligible, or to balance. The equilibrium equations express the condition that the total rate of arrival in any one state equals the total rate of departure; it is necessary here to treat the various substates separately. The range of temperatures  $\theta$  and  $T$  over which our approximate formulae hold depends on whether  $E_n^r/k\theta \gtrsim 1$  etc. and it is convenient first to specify the temperature conditions. Now

$$E_n^r/k\theta = 1.58 \times 10^5 (n^{-2} - r^{-2})/\theta.$$

This expression exceeds  $1.58 \times 10^5 (1/9)/\theta$  for all transitions to be used, and so exceeds unity for  $\theta > 1.75 \times 10^4$  °K. For  $\theta$  and  $T$  less than  $1.75 \times 10^4$  °K. we obtain the following equations.

(1) *The ionized state*

$$\begin{aligned} & 4.4 \times 10^{-10} N_e \theta^{\frac{1}{2}} \sum_1^3 N_n n^4 \exp(-E_n^i/k\theta) + 7.85 \times 10^9 \kappa \sum_1^3 \frac{N_n}{n^5} \{-Ei(-E_n^i/kT)\} \\ & \quad \text{collision ionization} \qquad \qquad \qquad \text{photoelectric ionization} \\ & = (3.28 \times 10^{-6} N_+ N_e / \theta^{3/2}) \sum_1^3 n^{-3} \exp(E_n^i/k\theta) \{-Ei(-E_n^i/k\theta)\} \\ & \quad \text{spontaneous recombination} \\ & \quad + 1.29 \times 10^{-22} N_e^2 N_+ / \theta. \\ & \quad \text{three-body recombination} \end{aligned}$$

(2) *2P*

$$\begin{aligned} & \frac{N_1 \tilde{\omega}_{2P} \kappa A_{2P,1S}}{\tilde{\omega}_{1S} \exp E_1^2/kT} + N_{3S} A_{3S,2P} + N_{3D} A_{3D,2P} \\ & \quad \text{radiative excitation} \quad \text{spontaneous de-excitation} \\ & \quad \text{from } n=1 \qquad \qquad \text{from } n=3 \\ & + \frac{8.0 \times 10^{-13} N_e}{\theta^{\frac{1}{2}}} \left( N_{3S} + \frac{N_{3P}}{3} + \frac{N_{3D}}{5} \right) 3 A_{3,2} + \frac{6.9 \times 10^{-15} N_1 N_e \cdot 4 A_{2,1} \exp(-E_1^2/k\theta)}{2(1-0.25)^2 \theta^{\frac{1}{2}}} \\ & \quad \text{superelastic collision from } n=3 \qquad \qquad \text{collision excitation from } n=1 \\ & \quad + \frac{3.28 \times 10^{-6} \tilde{\omega}_{2P} N_+ N_e \exp(E_2^i/k\theta) \{-Ei(-E_2^i/k\theta)\}}{2^6 \theta^{3/2}} \\ & \quad \text{spontaneous recombination} \\ & = N_{2P} A_{2P,1S} + \frac{N_{2P} \tilde{\omega}_{3S} \kappa A_{3S,2P}}{\tilde{\omega}_{2P} \exp(E_2^3/kT)} + \frac{N_{2P} \tilde{\omega}_{3D} \kappa A_{3D,2P}}{\tilde{\omega}_{2P} \exp(E_2^3/kT)} \\ & \quad \text{spontaneous} \qquad \qquad \text{radiative excitation to } n=3 \\ & \quad \text{de-excitation} \\ & + \frac{6.9 \times 10^{-15} N_{2P} N_e \cdot 9 A_{3,2} \exp(-E_2^3/k\theta)}{4(0.250-0.111)^2 \theta^{\frac{1}{2}}} + 7.85 \times 10^9 \kappa N_{2P} \{-Ei(-E_2^i/kT)\} 2^{-5} \\ & \quad \text{collision excitation to } n=3 \qquad \qquad \text{photoelectric ionization} \\ & \quad + 4.4 \times 10^{-10} N_e N_{2P} \cdot 2^4 \theta^{\frac{1}{2}} \exp(-E_2^i/k\theta). \\ & \quad \text{collision ionization} \end{aligned}$$

(3) 2S

$$\begin{aligned}
& \frac{6 \cdot 9 \times 10^{-15} N_1 N_e 4 A_{2,1} \exp(-E_1^2/k\theta)}{2(1-0 \cdot 25)^2 \theta^{\frac{1}{2}}} + \frac{2 \cdot 07 \times 10^{-11} N_e N_+ \tilde{\omega}_{2S}}{2^4 \theta^{\frac{1}{2}}} + N_{3P} A_{3P,2S} \\
& \quad \text{collision excitation from } n=1 \qquad \qquad \text{spontaneous recombination} \qquad \text{spontaneous de-excitation from } n=3 \\
& + \frac{8 \cdot 0 \times 10^{-13} N_e}{\theta^{\frac{1}{2}}} \left( N_{3S} + \frac{N_{3P}}{3} + \frac{N_{3D}}{5} \right) A_{3,2} \\
& \quad \text{superelastic collision from } n=3 \\
& = 7 \cdot 85 \times 10^9 \alpha N_{2S} \cdot 2^{-5} \{ -Ei(-E_2^i/kT) \} \\
& \quad \text{photoelectric ionization} \\
& + 4 \cdot 4 \times 10^{-10} N_e N_{2S} 2^4 \theta^{\frac{1}{2}} \exp(-E_2^i/k\theta) + \frac{N_{2S} \tilde{\omega}_{3P} \alpha A_{3P,2S}}{\tilde{\omega}_{2S} \exp[E_2^3/kT]} \\
& \quad \text{collision ionization} \qquad \qquad \text{radiative excitation to } n=3 \\
& + \frac{6 \cdot 9 \times 10^{-15} N_{2S} N_e 9 A_{3,2} \exp(-E_2^3/k\theta)}{4(0 \cdot 250 - 0 \cdot 111)^2 \theta^{\frac{1}{2}}} + 1 \cdot 16 \times 10^{-5} N_e N_{2S} / \theta^{\frac{1}{2}}. \\
& \quad \text{collision excitation to } n=3 \qquad \qquad \text{superelastic collision to } n=1
\end{aligned}$$

(4) 3S

$$\begin{aligned}
& \frac{6 \cdot 9 \times 10^{-15} N_1 N_e 9 A_{3,1} \exp(-E_1^3/k\theta)}{3(1-0 \cdot 111)^2 \theta^{\frac{1}{2}}} \\
& \quad \text{collision excitation from } n=1 \\
& + \frac{6 \cdot 9 \times 10^{-15} (N_{2S} + N_{2P}) N_e 9 A_{3,2} \exp(-E_2^3/k\theta)}{3 \times 4(0 \cdot 250 - 0 \cdot 111)^2 \theta^{\frac{1}{2}}} \\
& \quad \text{collision excitation from } n=2 \\
& + \frac{N_{2P} \tilde{\omega}_{3S} \alpha A_{3S,2P}}{\tilde{\omega}_{2P} \exp[E_2^3/kT]} + \frac{\tilde{\omega}_{3S} 3 \cdot 28 \times 10^{-6} N_+ N_e \{ -Ei(-E_3^i/k\theta) \} \exp(E_3^i/k\theta)}{2 \times 9 \times 27 \theta^{3/2}} \\
& \quad \text{radiative excitation from } n=2 \qquad \qquad \text{spontaneous recombination} \\
& + 1 \cdot 43 \times 10^{-23} N_e^2 N_+ / \theta \\
& \quad \text{three-body recombination} \\
& = \frac{7 \cdot 85 \times 10^9 \alpha}{3^5} N_{3S} \{ -Ei(-E_3^i/kT) \} + 4 \cdot 4 \times 10^{-10} N_e N_{3S} \cdot 3^4 \theta^{\frac{1}{2}} \exp(-E_3^i/k\theta) \\
& \quad \text{photoelectric ionization} \qquad \qquad \text{collision ionization} \\
& + N_{3S} A_{3S,2P} + 8 \cdot 0 \times 10^{-13} N_e N_{3S} \cdot 4 A_{3,2} / \theta^{\frac{1}{2}}. \\
& \quad \text{spontaneous de-excitation} \qquad \text{superelastic collision to } n=2 \\
& \quad \text{to } n=2
\end{aligned}$$

(5) 3P

$$\begin{aligned}
& \frac{6 \cdot 9 \times 10^{-15} N_1 N_e 9 A_{3,1} \exp(-E_1^3/k\theta)}{3(1-0 \cdot 111)^{2\theta^{\frac{1}{2}}}} + \frac{6 \cdot 9 \times 10^{-15} (N_{2S} + N_{2P}) N_e 9 A_{3,2} \exp(-E_2^3/k\theta)}{3 \times 4(0 \cdot 25 - 0 \cdot 111)^{2\theta^{\frac{1}{2}}}} \\
& \quad \text{collision excitation from } n=1 \qquad \qquad \text{collision excitation from } n=2 \\
& + \frac{N_1 \tilde{\omega}_{3P} \kappa A_{3P,1S}}{\tilde{\omega}_{1S} \exp(E_1^3/kT)} + \frac{N_{2S} \tilde{\omega}_{3P} \kappa A_{3P,2S}}{\tilde{\omega}_{2S} \exp(E_2^3/kT)} \\
& \quad \text{radiative excitation} \quad \text{radiative excitation} \\
& \quad \text{from } n=1 \qquad \qquad \text{from } n=2 \\
& + \frac{\tilde{\omega}_{3P} 3 \cdot 28 \times 10^{-6} N_+ N_e \{-Ei(-E_3^3/k\theta)\} \exp(E_3^3/k\theta)}{2 \times 9 \times 27\theta^{3/2}} + 4 \cdot 3 \times 10^{-23} N_e^2 N_+ / \theta \\
& \quad \text{spontaneous recombination} \qquad \qquad \text{three-body recombination} \\
& = \frac{7 \cdot 85 \times 10^9 \kappa N_{3P} \{-Ei(-E_3^3/kT)\}}{3^5} + 4 \cdot 4 \times 10^{-10} N_e N_{3P} 3^4 \theta^{\frac{1}{2}} \exp(-E_3^3/k\theta) \\
& \quad \text{photoelectric ionization} \qquad \qquad \text{collision ionization} \\
& + N_{3P} A_{3P,2S} + N_{3P} A_{3P,1S} + 8 \cdot 0 \times 10^{-13} N_e N_{3P} 4 A_{3,2} / 3\theta^{\frac{1}{2}}. \\
& \quad \text{spontaneous} \quad \text{spontaneous} \quad \text{superelastic collision to } n=2 \\
& \quad \text{de-excitation} \quad \text{de-excitation} \\
& \quad \text{to } n=2 \qquad \qquad \text{to } n=1
\end{aligned}$$

(6) 3D

$$\begin{aligned}
& \frac{6 \cdot 9 \times 10^{-15} N_1 N_e 9 A_{3,1} \exp(-E_1^3/k\theta)}{3(1-0 \cdot 111)^{2\theta^{\frac{1}{2}}}} + \frac{6 \cdot 9 \times 10^{-15} (N_{2S} + N_{2P}) N_e 9 A_{3,2} \exp(-E_2^3/k\theta)}{3 \times 4(0 \cdot 25 - 0 \cdot 111)^{2\theta^{\frac{1}{2}}}} \\
& \quad \text{collision excitation from } n=1 \qquad \qquad \text{collision excitation from } n=2 \\
& + \frac{N_{2P} \tilde{\omega}_{3D} \kappa A_{3D,2P}}{\tilde{\omega}_{2P} \exp(E_2^3/kT)} + 7 \cdot 16 \times 10^{-23} N_e^2 N_+ / \theta \\
& \quad \text{radiative} \qquad \qquad \text{three-body} \\
& \quad \text{excitation} \qquad \qquad \text{recombination} \\
& \quad \text{from } n=2 \\
& + \frac{3 \cdot 28 \times 10^{-6} \tilde{\omega}_{3D} N_+ N_e \{-Ei(-E_3^3/k\theta)\} \exp(E_3^3/k\theta)}{2 \times 9 \times 27\theta^{3/2}} \\
& \quad \text{spontaneous recombination} \\
& = \frac{7 \cdot 85 \times 10^9 \kappa N_{3D} \{-Ei(-E_3^3/kT)\}}{3^5} + 4 \cdot 4 \times 10^{-10} N_e N_{3D} 3^4 \theta^{\frac{1}{2}} \exp(-E_3^3/k\theta) \\
& \quad \text{photoelectric ionization} \qquad \qquad \text{collision ionization} \\
& + N_{3D} A_{3D,2P} + 8 \cdot 0 \times 10^{-13} N_e N_{3D} \cdot 4 A_{3,2} / 5\theta^{\frac{1}{2}}. \\
& \quad \text{spontaneous} \quad \text{superelastic collision} \\
& \quad \text{de-excitation} \quad \text{to } n=2 \\
& \quad \text{to } n=2
\end{aligned}$$



The solution of the above equations has for convenience been carried out for an electrically neutral atmosphere of pure hydrogen so that  $N_+ = N_e$ . On allocating numerical values to  $N_e$ ,  $T$ ,  $\theta$ , and  $\alpha$ , the corresponding values of  $N_1$ ,  $N_{2S}$ ,  $N_{2P}$ ,  $N_{3S}$ ,  $N_{3P}$ , and  $N_{3D}$  may then be obtained.

For  $\theta > 1.75 \times 10^4$  °K., it is necessary to make appropriate modifications to some of the terms in the above equations. These modifications are straightforward, and will not be listed here, although they have been included in all computations leading to the results tabulated below.

The computed populations are given in Tables 1 to 6 for  $T = 5 \times 10^3$  °K.,  $\alpha = 1, 0.5$  and  $0$ ,  $N_e = 10^{13}, 10^{12}, 10^{11}, 10^{10}, 10^9$  and, in some cases,  $10^8$  per cc., and  $\theta = 7.5 \times 10^3, 1.0 \times 10^4, 1.5 \times 10^4, 2.5 \times 10^4, 5.0 \times 10^4$  and  $1.0 \times 10^5$  °K. The computations have been carried out to an accuracy of about 1 per cent.

For comparison purposes, we may note that in thermodynamic equilibrium at a temperature  $\zeta$  the corresponding concentrations are given by

$$N_j = \frac{\tilde{\omega}_j N_+ N_e h^3 \exp(E_j^i/k\zeta)}{\tilde{\omega}_{1S}(2\pi mk\zeta)^{3/2}} = 4.2 \times 10^{-16} \tilde{\omega}_j N_e^2 \{\exp(E_j^i/k\zeta)\} / \tilde{\omega}_{1S} \zeta^{3/2}$$

in an atmosphere where  $N_+ = N_e$ . These values are given in the tables for  $N_e = 10^{13}$  per cc.,  $\zeta = T$ , and  $\zeta = \theta$ .

#### IV. DISCUSSION

Three general features of the above results may be noted here.

(i) For all the electron temperatures and for any given electron concentration, the population of the ground state (and the concentration of the neutral atoms as a whole) is less than in thermodynamic equilibrium with the electrons at the radiation temperature,  $T = 5 \times 10^3$  °K. In most cases this difference amounts to several powers of ten.

(ii) The population of the metastable  $2S$  state is much greater than that of the  $2P$  state, whereas in thermodynamic equilibrium the latter exceeds the former in the ratio of the statistical weights, i.e. 3 to 1.

(iii) While the 3-quantum atoms are far from being distributed among the substates in accordance with their statistical weights, the populations of the  $2P$  and 3-quantum states for the higher temperatures (say  $\theta > 1.5 \times 10^4$  °K.) are closer to those in thermodynamic equilibrium with the electrons at temperature  $\theta$  than at temperature  $T$ .

We now estimate the magnitude of the errors involved in the above results. The solutions of the equations have mostly been carried out to an accuracy of 1 per cent. or better, though the errors in  $N_{2P}$  and  $N_{3P}$  may in a few cases be slightly greater. The major sources of error are, however, in the equations themselves.

(a) The collision cross sections and the rate of ionization by radiation are uncertain to within a factor of the order of 2. This is of little consequence if the equilibrium population of a state is maintained by one excitation process and its reverse de-excitation process, for then the population of this state will be related to that of the other state in a manner independent of uncertainties in the

transition rates. For example, this applies to the  $2P$  state for  $\theta=7.5 \times 10^3$  and  $1.0 \times 10^4$  °K.,  $N_e=10^8$  to  $10^{10}$  per cc., arrivals in the  $2P$  state being due mainly to absorption of radiation in the ground state, and departures from the  $2P$  state being mainly spontaneous transitions to the ground state. Hence in this case a simple approximation holds, viz.

$$N_{2P} = \alpha \tilde{\omega}_{2P} N_1 \{ \exp(-E_1^2/kT) \} / \tilde{\omega}_1.$$

In most cases, however, such conditions do not apply, and the results themselves will be uncertain to within a factor of the order of 2.

(b) It has already been pointed out that elastic collisions causing  $2S \rightarrow 2P$  transitions have been neglected, and that if the product of the electron velocity and collision cross section be constant and equal to  $10^{-8}$  cm.<sup>3</sup> per second, say, then there will be  $10^{-8} N_{2S} N_e$  such transitions per second. We may compare this with the rate of transition from the  $2S$  state by all other types of electron collisions, which ranges from  $1.8 \times 10^{-7} N_{2S} N_e$  for  $\theta=7.5 \times 10^3$  °K. to  $8.4 \times 10^{-7} N_{2S} N_e$  for  $\theta=1.0 \times 10^5$  °K. Also there is a total of  $8.5 \times 10^5 N_{2S} \alpha$  transitions from the  $2S$  to the  $3P$  and ionized states due to the absorption of dilute radiation from a black body at a temperature of  $5 \times 10^3$  °K. Thus we see that unless the cross section for collisions producing the  $2S \rightarrow 2P$  transition is abnormally great, such collisions are not likely to be important. This is particularly so when  $N_e$  is small, say less than  $10^{11}$  per cc.

While we may feel justified for the present in neglecting these collisions, it is clear that their importance still remains an open question, which will be answered only after a reliable estimate of the cross section becomes available.

(c) A further source of major error arises from the neglect of transitions to and from the states  $n \geq 4$ . We may here consider separately transitions involving radiation and transitions involving only collisions. It may be noted that the populations of the various states tend towards those in thermodynamic equilibrium with the electrons as the principal quantum number increases. By inspection of the tables, we can readily see that the ratio of the population of a higher state to that of a lower state is greater than corresponds to thermodynamic equilibrium at a temperature  $T$ , except where the lower state is the metastable  $2S$  state. Thus, considering only transitions involving radiation, we see that generally more of these occur from a higher to a lower state than vice versa. We over-estimate the error made in neglecting such transitions if we consider only spontaneous emissions. An approximation to these is obtained by assuming for  $n \geq 4$  that the populations are as in thermodynamic equilibrium with the electrons, so that

$$N_j \approx 4.2 \times 10^{-16} (\tilde{\omega}_j / \tilde{\omega}_1) (N_e^2 / \theta^{3/2}) \exp(E_j^i / k\theta).$$

The most serious error comes from a neglect of the 4-quantum state. For this

$$N_j \approx (\tilde{\omega}_j / \tilde{\omega}_1) N_e^2 x,$$

where  $x$  ranges from  $2.4 \times 10^{-21}$  for  $\theta=7.5 \times 10^3$  °K. to  $1.5 \times 10^{-23}$  for  $\theta=1.0 \times 10^5$  °K. We then find, for example, that the rate of  $4D \rightarrow 2P$  transition is  $N_{4D} A_{4D,2P}$ , and ranges from  $2.5 \times 10^{-13} N_e^2$  to  $1.5 \times 10^{-15} N_e^2$  for the above

temperatures. These values are small in comparison with the total rates of transition to the  $2P$  state. Similarly, we find that in most cases transition rates from the 4-quantum states to a given lower state are considerably less than 10 per cent. of the total rates of transition to this state; in a few cases involving the  $3D$  state, the rate of  $4F \rightarrow 3D$  transition approaches 30 per cent. of the total. The neglect of emissions from the 4-quantum state is thus not as serious as the uncertainty in the transition probabilities discussed in (a).

For the  $2S$  state, we need consider only excitations to the  $4P$  state. The rate of these is  $5.6 \times 10^4 \times N_{2S}$ , which is less than one-tenth of the rate of other departures from the  $2S$  state.

In considering transitions to and from the upper states due to electron collision, we may note that the ratio of the population of a given lower state to that of a given upper state is higher than in thermodynamic equilibrium at the electron temperature,  $\theta$ , so that we over-estimate the error if we neglect super-elastic collisions. The most serious error comes from a neglect of transitions to the 4-quantum state. The rate of collision excitation from 2- to 4-quantum states ranges from  $1.5 \times 10^{-9} N_2 N_e$  for  $\theta = 7.5 \times 10^3$  °K. to  $1.5 \times 10^{-8} N_2 N_e$  for  $\theta = 1.0 \times 10^5$  °K.; the rate of excitation from 3- to 4-quantum states ranges from  $4.4 \times 10^{-7} N_3 N_e$  for  $\theta = 7.5 \times 10^3$  °K. to  $3.1 \times 10^{-7} N_3 N_e$  for  $\theta = 1.0 \times 10^5$  °K. Except for transitions from the  $N_{3S}$  state when  $N_e = 10^{13}$ , these collision excitations represent only a small fraction of the total transitions from any state. In the case of the  $3S$  state when  $N_e = 10^{13}$  per cc., such transitions occur at approximately the same rate as all other transitions from the  $3S$  state.

We are thus justified in stating that neglect of transitions to and from states  $n \geq 4$  has not introduced errors more serious than those due to uncertainties in the transition rates themselves.

(d) Excitation and de-excitation by neutral atom-neutral atom and neutral atom-ion collisions have been tacitly neglected. For atoms having energies of the magnitudes considered in this paper the cross sections for such collisions are very much smaller than for excitation by electron collisions, so that when the electron concentration is not too small neglect of the above collisions leads to no appreciable error. But when, as in several cases listed in Table 1, the neutral atom concentration much exceeds that of the electrons, neutral atom-neutral atom collisions may be significant. However, the results obtained here for temperatures of  $10^4$  °K. or higher are free from such error.

In general, the values given in Tables 2-6 are correct to within a factor of the order of 2, provided the effect of elastic collisions  $2S \rightarrow 2P$  is small.

## V. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Physics, C.S.I.R.

## VI. REFERENCES

- (1) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1**: 275 (1948).
- (2) SLACK, F. G.—*Phys. Rev.* **31**: 527 (1928).
- (3) FOWLER, R. H.—“Statistical Mechanics”, p. 478. (Cambridge University Press, 1929.)
- (4) MASSEY, H. S. W., and MOHR, C. B. O.—*Proc. Roy. Soc. A* **132**: 605 (1931).

TABLE 1

$$T=5 \times 10^3 \quad \theta=7.5 \times 10^3$$

Conditions	$N_e$	$N_1$	$N_{2S}$	$N_{2P}$	$N_{3S}$	$N_{3P}$	$N_{3D}$
Therm. Equilib. Values at $5 \times 10^3$ °K. . .	$10^{13}$	$6.3 \times 10^{18}$	$3.2 \times 10^8$	$9.7 \times 10^8$	$4.0 \times 10^6$	$1.2 \times 10^7$	$2.0 \times 10^7$
Therm. Equilib. Values at $7.5 \times 10^3$ °K.	$10^{13}$	$3.3 \times 10^{13}$	$1.2 \times 10^7$	$3.8 \times 10^7$	$6.7 \times 10^5$	$2.0 \times 10^6$	$3.4 \times 10^6$
$x=1$	$10^{13}$	$9.3 \times 10^{15}$	$7.6 \times 10^8$	$4.4 \times 10^8$	$2.7 \times 10^6$	$3.5 \times 10^6$	$1.2 \times 10^6$
0.5		$9.6 \times 10^{15}$	$9.3 \times 10^8$	$3.8 \times 10^8$	$3.2 \times 10^6$	$2.4 \times 10^6$	$1.3 \times 10^6$
0		$1.0 \times 10^{16}$	$1.2 \times 10^9$	$3.2 \times 10^8$	$4.0 \times 10^6$	$5.4 \times 10^5$	$1.6 \times 10^6$
1	$10^{12}$	$1.1 \times 10^{15}$	$2.2 \times 10^7$	$2.1 \times 10^5$	$3.7 \times 10^4$	$9.9 \times 10^4$	$9.6 \times 10^3$
0.5		$1.1 \times 10^{15}$	$3.8 \times 10^7$	$1.2 \times 10^5$	$5.2 \times 10^4$	$8.7 \times 10^4$	$8.3 \times 10^3$
0		$9.8 \times 10^{14}$	$1.2 \times 10^8$	$3.2 \times 10^4$	$1.2 \times 10^5$	$5.2 \times 10^3$	$1.6 \times 10^4$
1	$10^{11}$	$1.0 \times 10^{14}$	$2.4 \times 10^5$	$1.5 \times 10^4$	$2.8 \times 10^2$	$1.2 \times 10^3$	$3.4 \times 10^2$
0.5		$1.2 \times 10^{14}$	$5.5 \times 10^5$	$9.2 \times 10^3$	$2.9 \times 10^2$	$1.3 \times 10^3$	$1.3 \times 10^2$
0		$1.0 \times 10^{14}$	$1.2 \times 10^7$	$3.4 \times 10^2$	$1.6 \times 10^3$	$5.5 \times 10$	$1.6 \times 10^2$
1	$10^{10}$	$3.6 \times 10^{12}$	$1.1 \times 10^3$	$5.5 \times 10^2$	3.0	$1.1 \times 10$	$1.1 \times 10$
0.5		$6.4 \times 10^{12}$	$3.5 \times 10^3$	$4.9 \times 10^2$	2.4	$1.3 \times 10$	5.2
0		$1.0 \times 10^{14}$	$1.2 \times 10^6$	$3.4 \times 10^2$	$1.6 \times 10$	$5.5 \times 10^{-1}$	1.7
1	$10^9$	$4.8 \times 10^{10}$	3.7	7.4	$3.2 \times 10^{-2}$	$9.7 \times 10^{-2}$	$1.5 \times 10^{-1}$
0.5		$1.2 \times 10^{11}$	$1.2 \times 10$	8.9	$2.1 \times 10^{-2}$	$1.2 \times 10^{-1}$	$9.2 \times 10^{-2}$
0		$1.0 \times 10^{14}$	$1.2 \times 10^5$	$3.4 \times 10^2$	$1.7 \times 10^{-1}$	$5.5 \times 10^{-3}$	$1.7 \times 10^{-2}$
1	$10^8$	$5.0 \times 10^8$	$2.7 \times 10^{-2}$	$7.7 \times 10^{-2}$	$3.2 \times 10^{-4}$	$9.6 \times 10^{-4}$	$1.6 \times 10^{-3}$
0.5		$1.3 \times 10^9$	$7.0 \times 10^{-2}$	$9.7 \times 10^{-2}$	$2.1 \times 10^{-4}$	$1.2 \times 10^{-3}$	$1.0 \times 10^{-3}$
0		$1.0 \times 10^{14}$	$1.2 \times 10^4$	$3.4 \times 10^2$	$1.7 \times 10^{-3}$	$5.5 \times 10^{-3}$	$1.7 \times 10^{-4}$



TABLE 2

$$T = 5 \times 10^3 \quad \theta = 1.0 \times 10^4$$

Conditions	$N_e$	$N_1$	$N_{2S}$	$N_{2P}$	$N_{3S}$	$N_{3P}$	$N_{3D}$
Therm. Equilib. Values at $1.0 \times 10^4$ °K.	$10^{13}$	$3.0 \times 10^{11}$	$2.2 \times 10^6$	$4.4 \times 10^5$	$2.4 \times 10^5$	$7.3 \times 10^5$	$1.2 \times 10^6$
$\kappa=1$	$10^{13}$	$5.2 \times 10^{13}$	$1.8 \times 10^8$	$7.9 \times 10^5$	$1.2 \times 10^6$	$9.2 \times 10^5$	$5.5 \times 10^5$
0.5		$4.9 \times 10^{13}$	$2.0 \times 10^8$	$7.2 \times 10^5$	$1.3 \times 10^6$	$6.1 \times 10^5$	$5.8 \times 10^5$
0		$4.5 \times 10^{13}$	$2.3 \times 10^8$	$7.0 \times 10^5$	$1.4 \times 10^6$	$2.0 \times 10^5$	$6.1 \times 10^5$
1	$10^{12}$	$1.0 \times 10^{13}$	$9.2 \times 10^6$	$1.8 \times 10^4$	$2.9 \times 10^4$	$4.1 \times 10^4$	$4.4 \times 10^3$
0.5		$8.6 \times 10^{12}$	$1.3 \times 10^7$	$1.3 \times 10^4$	$3.3 \times 10^4$	$3.0 \times 10^4$	$4.8 \times 10^3$
0		$4.4 \times 10^{12}$	$2.2 \times 10^7$	$7.0 \times 10^3$	$4.2 \times 10^4$	$1.9 \times 10^3$	$5.7 \times 10^3$
1	$10^{11}$	$1.6 \times 10^{12}$	$1.7 \times 10^5$	$4.5 \times 10^2$	$3.1 \times 10^2$	$7.5 \times 10^2$	$4.5 \times 10$
0.5		$1.5 \times 10^{12}$	$3.2 \times 10^5$	$3.3 \times 10^2$	$3.4 \times 10^2$	$7.2 \times 10^2$	$4.2 \times 10$
0		$4.6 \times 10^{11}$	$2.3 \times 10^6$	$7.3 \times 10$	$5.5 \times 10^2$	$1.9 \times 10$	$5.9 \times 10$
1	$10^{10}$	$1.6 \times 10^{11}$	$1.8 \times 10^3$	$2.6 \times 10$	3.0	8.1	$8.8 \times 10^{-1}$
0.5		$1.7 \times 10^{11}$	$3.8 \times 10^3$	$1.5 \times 10$	3.3	8.4	$5.2 \times 10^{-1}$
0		$4.6 \times 10^{10}$	$2.3 \times 10^5$	$7.3 \times 10^{-1}$	5.6	$2.0 \times 10^{-1}$	$5.9 \times 10^{-1}$
1	$10^9$	$1.2 \times 10^{10}$	$1.4 \times 10$	1.8	$2.9 \times 10^{-2}$	$8.1 \times 10^{-2}$	$4.0 \times 10^{-2}$
0.5		$1.5 \times 10^{10}$	$3.4 \times 10$	1.2	$3.0 \times 10^{-2}$	$8.8 \times 10^{-2}$	$1.5 \times 10^{-2}$
0		$4.6 \times 10^9$	$2.3 \times 10^4$	$7.4 \times 10^{-3}$	$5.8 \times 10^{-2}$	$2.0 \times 10^{-3}$	$6.0 \times 10^{-3}$
1	$10^8$	$3.3 \times 10^8$	$5.4 \times 10^{-2}$	$5.1 \times 10^{-2}$	$2.7 \times 10^{-4}$	$7.9 \times 10^{-4}$	$1.0 \times 10^{-3}$
0.5		$6.5 \times 10^8$	$1.8 \times 10^{-1}$	$5.0 \times 10^{-2}$	$2.3 \times 10^{-4}$	$9.4 \times 10^{-4}$	$5.3 \times 10^{-4}$
0		$4.6 \times 10^8$	$2.3 \times 10^3$	$7.4 \times 10^{-5}$	$5.8 \times 10^{-4}$	$2.0 \times 10^{-5}$	$6.0 \times 10^{-5}$



TABLE 3  
 $T=5 \times 10^3$      $\theta=1.5 \times 10^4$

Conditions	$N_e$	$N_1$	$N_{2S}$	$N_{2P}$	$N_{3S}$	$N_{3P}$	$N_{3D}$
Therm. Equilib. Values at $1.5 \times 10^4$ °K.	$10^{13}$	$8.3 \times 10^8$	$3.1 \times 10^5$	$9.4 \times 10^5$	$7.4 \times 10^4$	$2.2 \times 10^5$	$3.7 \times 10^5$
$\alpha=1$	$10^{13}$	$2.7 \times 10^{11}$	$3.3 \times 10^7$	$1.9 \times 10^5$	$3.7 \times 10^5$	$2.0 \times 10^5$	$2.1 \times 10^5$
0.5		$2.5 \times 10^{11}$	$3.4 \times 10^7$	$1.8 \times 10^5$	$3.8 \times 10^5$	$1.4 \times 10^5$	$2.2 \times 10^5$
0		$2.2 \times 10^{11}$	$3.7 \times 10^7$	$1.6 \times 10^5$	$3.8 \times 10^5$	$7.1 \times 10^4$	$2.2 \times 10^5$
1	$10^{12}$	$5.8 \times 10^{10}$	$2.1 \times 10^6$	$3.7 \times 10^3$	$1.3 \times 10^4$	$9.9 \times 10^3$	$2.1 \times 10^3$
0.5		$4.5 \times 10^{10}$	$2.7 \times 10^6$	$2.9 \times 10^3$	$1.3 \times 10^4$	$6.4 \times 10^3$	$2.0 \times 10^3$
0		$2.8 \times 10^{10}$	$3.6 \times 10^6$	$1.7 \times 10^3$	$1.2 \times 10^4$	$6.2 \times 10^3$	$2.0 \times 10^3$
1	$10^{11}$	$1.2 \times 10^{10}$	$5.4 \times 10^4$	$7.2 \times 10$	$2.0 \times 10^2$	$2.4 \times 10^2$	$2.3 \times 10$
0.5		$1.0 \times 10^{10}$	$9.3 \times 10^4$	$6.4 \times 10$	$2.0 \times 10^2$	$2.1 \times 10^2$	$2.3 \times 10$
0		$2.3 \times 10^9$	$3.6 \times 10^5$	$1.7 \times 10$	$1.7 \times 10^2$	6.3	$2.0 \times 10$
1	$10^{10}$	$1.3 \times 10^9$	$6.1 \times 10^2$	$9.8 \times 10^{-1}$	2.1	2.7	$2.4 \times 10^{-1}$
0.5		$1.3 \times 10^9$	$1.2 \times 10^3$	$8.8 \times 10^{-1}$	2.2	2.7	$2.4 \times 10^{-1}$
0		$2.4 \times 10^8$	$3.8 \times 10^4$	$1.8 \times 10^{-1}$	1.8	$6.3 \times 10^{-2}$	$2.0 \times 10^{-1}$
1	$10^9$	$1.3 \times 10^8$	6.2	$2.8 \times 10^{-2}$	$2.2 \times 10^{-2}$	$2.8 \times 10^{-2}$	$3.0 \times 10^{-3}$
0.5		$1.3 \times 10^8$	$1.3 \times 10$	$1.8 \times 10^{-2}$	$2.2 \times 10^{-2}$	$2.8 \times 10^{-2}$	$2.6 \times 10^{-3}$
0		$2.3 \times 10^8$	$3.8 \times 10^4$	$1.8 \times 10^{-3}$	$1.8 \times 10^{-2}$	$6.3 \times 10^{-4}$	$2.0 \times 10^{-3}$

TABLE 4

$$T=5 \times 10^3 \quad \theta=2.5 \times 10^4$$

Conditions	$N_e$	$N_1$	$N_{2S}$	$N_{2P}$	$N_{3S}$	$N_{3P}$	$N_{3D}$
Therm. Equilib. Values at $2.5 \times 10^4$ °K.	$10^{13}$	$5.9 \times 10^6$	$5.2 \times 10^4$	$1.5 \times 10^5$	$2.2 \times 10^4$	$6.5 \times 10^4$	$1.1 \times 10^5$
$\kappa=1$	$10^{13}$	$4.0 \times 10^9$	$5.8 \times 10^6$	$5.6 \times 10^4$	$1.4 \times 10^5$	$4.8 \times 10^4$	$7.9 \times 10^4$
0.5		$3.8 \times 10^9$	$6.0 \times 10^6$	$5.4 \times 10^4$	$1.4 \times 10^5$	$3.7 \times 10^4$	$7.9 \times 10^4$
0		$3.6 \times 10^9$	$6.2 \times 10^6$	$5.2 \times 10^4$	$1.4 \times 10^5$	$2.5 \times 10^4$	$7.9 \times 10^4$
1	$10^{12}$	$7.1 \times 10^8$	$4.1 \times 10^5$	$9.0 \times 10^2$	$4.7 \times 10^3$	$2.0 \times 10^3$	$8.0 \times 10^2$
0.5		$5.8 \times 10^8$	$5.0 \times 10^5$	$7.7 \times 10^2$	$4.6 \times 10^3$	$1.3 \times 10^3$	$7.8 \times 10^2$
0		$3.7 \times 10^8$	$6.4 \times 10^5$	$5.5 \times 10^2$	$4.4 \times 10^3$	$2.3 \times 10^2$	$7.5 \times 10^2$
1	$10^{11}$	$1.3 \times 10^8$	$1.0 \times 10^4$	$1.5 \times 10$	$7.2 \times 10$	$4.9 \times 10$	9.1
0.5		$1.1 \times 10^8$	$1.8 \times 10^4$	$1.4 \times 10$	$7.2 \times 10$	$4.2 \times 10$	9.0
0		$3.8 \times 10^7$	$6.5 \times 10^4$	5.8	$5.9 \times 10$	2.3	7.6
1	$10^{10}$	$1.4 \times 10^7$	$1.2 \times 10^2$	$1.7 \times 10^{-1}$	$7.8 \times 10^{-1}$	$5.7 \times 10^{-1}$	$9.4 \times 10^{-2}$
0.5		$1.4 \times 10^7$	$2.5 \times 10^2$	$1.7 \times 10^{-1}$	$8.0 \times 10^{-1}$	$5.6 \times 10^{-1}$	$9.6 \times 10^{-2}$
0		$3.9 \times 10^6$	$6.6 \times 10^3$	$5.8 \times 10^{-2}$	$6.2 \times 10^{-1}$	$2.3 \times 10^{-2}$	$7.6 \times 10^{-2}$
1	$10^9$	$1.4 \times 10^6$	1.3	$1.9 \times 10^{-3}$	$7.9 \times 10^{-3}$	$5.8 \times 10^{-3}$	$9.5 \times 10^{-4}$
0.5		$1.5 \times 10^6$	2.5	$1.8 \times 10^{-3}$	$8.1 \times 10^{-3}$	$5.8 \times 10^{-3}$	$9.6 \times 10^{-4}$
0		$3.9 \times 10^5$	$6.6 \times 10^2$	$5.8 \times 10^{-4}$	$6.2 \times 10^{-3}$	$2.3 \times 10^{-4}$	$7.6 \times 10^{-4}$

TABLE 5

$$T=5 \times 10^3 \quad \theta=5 \times 10^4$$

Conditions	$N_e$	$N_1$	$N_{2S}$	$N_{2P}$	$N_{3S}$	$N_{3P}$	$N_{3D}$
Therm. Equilib. Values at $5 \times 10^4$ °K. . .	$10^{13}$	$8.8 \times 10^4$	$8.3 \times 10^3$	$2.5 \times 10^4$	$5.3 \times 10^3$	$1.6 \times 10^4$	$2.7 \times 10^4$
$\alpha=1$	$10^{13}$	$1.2 \times 10^8$	$1.0 \times 10^8$	$1.6 \times 10^4$	$4.0 \times 10^4$	$1.2 \times 10^4$	$2.7 \times 10^4$
0.5		$1.2 \times 10^8$	$1.0 \times 10^8$	$1.6 \times 10^4$	$4.0 \times 10^4$	$1.0 \times 10^4$	$2.7 \times 10^4$
0		$1.2 \times 10^8$	$1.1 \times 10^8$	$1.6 \times 10^4$	$4.1 \times 10^4$	$8.6 \times 10^3$	$2.7 \times 10^4$
1	$10^{12}$	$1.6 \times 10^7$	$6.9 \times 10^4$	$2.0 \times 10^2$	$1.4 \times 10^3$	$3.8 \times 10^2$	$2.7 \times 10^2$
0.5		$1.5 \times 10^7$	$8.4 \times 10^4$	$1.9 \times 10^2$	$1.4 \times 10^3$	$2.6 \times 10^2$	$2.8 \times 10^2$
0		$1.2 \times 10^7$	$1.1 \times 10^5$	$1.7 \times 10^2$	$1.4 \times 10^3$	$8.2 \times 10$	$2.8 \times 10^2$
1	$10^{11}$	$2.4 \times 10^6$	$1.7 \times 10^3$	2.6	$1.7 \times 10$	8.0	2.7
0.5		$2.2 \times 10^6$	$3.0 \times 10^3$	2.5	$1.8 \times 10$	7.3	2.7
0		$1.2 \times 10^6$	$1.1 \times 10^4$	1.7	$1.9 \times 10$	8.1	2.8
1	$10^{10}$	$2.6 \times 10^5$	$2.0 \times 10$	$2.7 \times 10^{-2}$	$1.8 \times 10^{-1}$	$9.5 \times 10^{-2}$	$2.6 \times 10^{-2}$
0.5		$2.6 \times 10^5$	$4.0 \times 10$	$2.7 \times 10^{-2}$	$1.8 \times 10^{-1}$	$9.4 \times 10^{-2}$	$2.7 \times 10^{-2}$
0		$1.2 \times 10^5$	$1.1 \times 10^3$	$1.7 \times 10^{-2}$	$2.1 \times 10^{-1}$	$8.3 \times 10^{-3}$	$2.9 \times 10^{-2}$
1	$10^9$	$2.6 \times 10^4$	$2.0 \times 10^{-1}$	$2.8 \times 10^{-4}$	$1.8 \times 10^{-3}$	$9.7 \times 10^{-4}$	$2.7 \times 10^{-4}$
0.5		$2.6 \times 10^4$	$4.1 \times 10^{-1}$	$2.8 \times 10^{-4}$	$1.8 \times 10^{-3}$	$9.7 \times 10^{-4}$	$2.7 \times 10^{-4}$
0		$1.2 \times 10^4$	$1.1 \times 10^2$	$1.7 \times 10^{-4}$	$2.1 \times 10^{-3}$	$8.2 \times 10^{-5}$	$2.9 \times 10^{-4}$

TABLE 6

$$T=5 \times 10^3 \quad \theta=1.0 \times 10^5$$

Conditions	$N_e$	$N_1$	$N_{2S}$	$N_{2P}$	$N_{3S}$	$N_{3P}$	$N_{3D}$
Therm. Equilib. Values at $1.0 \times 10^5$ °K.	$10^{13}$	$6.4 \times 10^3$	$2.0 \times 10^3$	$5.9 \times 10^3$	$1.6 \times 10^3$	$4.8 \times 10^3$	$7.9 \times 10^3$
$\kappa=1$	$10^{13}$	$1.5 \times 10^7$	$3.4 \times 10^5$	$5.7 \times 10^3$	$1.6 \times 10^4$	$4.6 \times 10^3$	$1.1 \times 10^4$
0.5		$1.4 \times 10^7$	$3.5 \times 10^5$	$5.6 \times 10^3$	$1.6 \times 10^4$	$4.0 \times 10^3$	$1.1 \times 10^4$
0		$1.4 \times 10^7$	$3.6 \times 10^5$	$5.6 \times 10^3$	$1.7 \times 10^4$	$3.4 \times 10^3$	$1.1 \times 10^4$
1	$10^{12}$	$1.8 \times 10^6$	$2.1 \times 10^4$	$6.4 \times 10$	$4.8 \times 10^2$	$1.2 \times 10^2$	$1.1 \times 10^2$
0.5		$1.7 \times 10^6$	$2.7 \times 10^4$	$6.3 \times 10$	$5.1 \times 10^2$	$9.0 \times 10$	$1.1 \times 10^2$
0		$1.5 \times 10^6$	$3.7 \times 10^4$	$6.1 \times 10$	$5.8 \times 10^2$	$3.4 \times 10$	$1.2 \times 10^2$
1	$10^{11}$	$2.1 \times 10^5$	$4.6 \times 10^2$	$6.9 \times 10^{-1}$	5.1	2.2	$9.5 \times 10^{-1}$
0.5		$2.1 \times 10^5$	$7.5 \times 10^2$	$6.9 \times 10^{-1}$	5.4	1.9	$9.8 \times 10^{-1}$
0		$1.5 \times 10^5$	$3.7 \times 10^3$	$6.3 \times 10^{-1}$	7.8	$3.4 \times 10^{-1}$	1.2
1	$10^{10}$	$2.2 \times 10^4$	5.2	$7.0 \times 10^{-3}$	$5.0 \times 10^{-2}$	$2.5 \times 10^{-2}$	$9.2 \times 10^{-3}$
0.5		$2.2 \times 10^4$	$1.0 \times 10$	$7.0 \times 10^{-3}$	$5.1 \times 10^{-2}$	$2.5 \times 10^{-2}$	$9.3 \times 10^{-3}$
0		$1.5 \times 10^4$	$3.7 \times 10^2$	$6.3 \times 10^{-3}$	$8.0 \times 10^{-2}$	$3.4 \times 10^{-3}$	$1.2 \times 10^{-2}$
1	$10^9$	$2.2 \times 10^3$	$5.2 \times 10^{-2}$	$7.2 \times 10^{-5}$	$4.9 \times 10^{-4}$	$2.5 \times 10^{-4}$	$9.3 \times 10^{-5}$
0.5		$2.2 \times 10^3$	$1.0 \times 10^{-1}$	$7.0 \times 10^{-5}$	$5.0 \times 10^{-4}$	$2.5 \times 10^{-4}$	$9.2 \times 10^{-5}$
0		$1.5 \times 10^3$	$3.7 \times 10$	$6.3 \times 10^{-5}$	$8.0 \times 10^{-4}$	$3.4 \times 10^{-5}$	$1.2 \times 10^{-4}$

# HYDROGEN ATMOSPHERES IN THE ABSENCE OF THERMODYNAMIC EQUILIBRIUM

## III. THE EMISSION OF RADIATION

By R. G. GIOVANELLI\*

[*Manuscript received June 29, 1948*]

### *Summary*

Well-known theories of the profile and intensity of spectral lines are first extended to lines involving a number of degenerate substates and under conditions where thermodynamic equilibrium does not apply. It is shown that much information can be obtained from an effective temperature and the absorption coefficient, and these two properties are then tabulated for the  $L\alpha$ ,  $L\beta$ , and  $H\alpha$  lines of hydrogen. The physical conditions considered cover electron concentrations from  $10^8$  to  $10^{13}$  per cc., gas temperatures from  $7.5 \times 10^3$  to  $1.0 \times 10^5$  °K., and incident black-body radiation (5000 °K.) diluted by a factor 1, 0.5, or 0.

## I. INTRODUCTION

In two previous papers(1, 2)† the populations of the upper and of the lower ( $n=1, 2$ , and 3) states and substates of neutral hydrogen atoms were determined as functions of  $N_e$ , the electron concentration, for a wide range of electron temperatures and in the presence of diluted black-body radiation at a temperature of 5000 °K. The populations of atoms in the upper states were shown to be the same as those in thermodynamic equilibrium with the electrons when  $n$ , the principal quantum number, is sufficiently large. The populations computed for the lower states involve considerable uncertainties; but these do not exceed a factor of the order of 2 if transitions  $2S \rightarrow 2P$  by electron collision are few in comparison with other transitions from the  $2S$  state, as was assumed in paper II.

The results of the two papers can be applied to the observed intensities of the spectral lines of the gas. Starting with the well-established theories of the profile and intensity of spectral lines for cases in which the emission or absorption is due to electron transitions between two non-degenerate substates alone, we investigate in the present paper the case of a spectral line due to emission or absorption from a number of degenerate substates. The results are applied to the  $H\alpha$ ,  $L\alpha$ , and  $L\beta$  lines of hydrogen for conditions where the electron concentration and temperature, and incident spectral radiation densities are known.

\* Division of Physics, C.S.I.R.

† These papers are subsequently referred to as papers I and II respectively.



This work is intended mainly for application in chromospheric research, and the temperatures and electron concentrations for which data are given are those which may be expected to occur somewhere in the sun's atmosphere. The theoretical discussion, however, is quite general and applies to any conditions.

## II. THE EMISSION OF RADIATION

In the absence of radiation,  $N_j A_{j,l}$  quanta are emitted per unit volume per second by atoms in the  $j$  state which make spontaneous transitions to the  $l$  state,  $A_{j,l}$  being the Einstein coefficient of spontaneous transition, and  $N_j$ , the concentration of atoms in the  $j$  state. In the presence of dilute black-body radiation at a temperature  $T$ , dilution factor  $\kappa$ , the corresponding number of quanta emitted is

$$N_j A_{j,l} \{1 + \kappa / [\exp(E_l^j / kT) - 1]\}.$$

Weisskopf and Wigner showed(3) that the energy emitted in a frequency interval  $d\nu$  of a spectral line by stationary atoms, in the absence of collision broadening, is distributed in a manner

$$J_\nu d\nu = \frac{(A_j + A_l) d\nu}{\frac{1}{4}(A_j + A_l)^2 + 4\pi^2(\nu - E_l^j/h)^2}$$

where  $J_\nu$  is normalized so that  $\int_0^\infty J_\nu d\nu = 1$ . In this expression,  $A_j$  and  $A_l$  are

the probabilities of transition from the two states  $j$  and  $l$ , i.e. they are the reciprocals of the mean lifetimes in these states. Weisskopf has pointed out(4) that a similar formula applies where the line is broadened by collisions if to  $A_j$  and  $A_l$  be added a term equal to the reciprocal of the mean time between collisions. Thus

$$J_\nu d\nu = \frac{A d\nu}{A^2/4 + 4\pi^2(\nu - E_l^j/h)^2},$$

where  $A$  will be not less than  $A_j + A_l$ . The energy emitted in a frequency interval  $d\nu$  by transitions of stationary atoms from the  $j$  to the  $l$  state is then given by

$$(I_\nu)_{j,l} d\nu = \frac{N_j A_{j,l} E_l^j \cdot A d\nu \{1 + \kappa / [\exp(E_l^j / kT) - 1]\}}{A^2/4 + 4\pi^2(\nu - \nu_0)^2},$$

where  $\nu_0 = E_l^j/h$ .

In general the emitting atoms will not be stationary but will have a distribution of velocities resulting, as is well known, in a Doppler broadening of the emission line. If these atom velocities are Maxwellian at a temperature  $\phi$ , the number  $dN_j$ , having line of sight velocities in the interval  $du$ , is given by

$$dN_j = N_j (m/2\pi k\phi)^{\frac{1}{2}} \exp(-mu^2/2k\phi) \cdot du.$$

Thus the energy emitted per unit volume per second into solid angle  $d\Omega$  near the line of sight is

$$(I_{\nu})_{j,l} d\nu d\Omega/4\pi$$

$$= (d\Omega/4\pi) \int_{-\infty}^{\infty} \frac{N_j(m/2\pi k\varphi)^{\frac{1}{2}} \exp(-mu^2/2k\varphi) \cdot A_{j,l} E_l^j \cdot A d\nu \{1 + \kappa/[\exp(-E_l^j/kT) - 1]\} du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2}$$

$$= \frac{d\Omega}{4\pi} \left( \frac{m}{2\pi k\varphi} \right)^{\frac{1}{2}} N_j A_{j,l} E_l^j A \cdot d\nu \{1 + \kappa/[\exp(-E_l^j/kT) - 1]\}$$

$$\times \int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) \cdot du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2}.$$

If there are several substates of an energy level from which quanta of the same frequency  $\nu$  may be emitted, the emission will be obtained by summation over all such energy levels, viz.

$$\frac{d\Omega}{4\pi} I_{\nu} d\nu = \frac{d\Omega}{4\pi} \sum_j (I_{\nu})_{j,l} d\nu.$$

The integral appearing in the above equations cannot be evaluated in finite terms. When  $\nu - \nu_0$  is small enough, we find that

$$\int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) \cdot du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2} \approx \int_{-\infty}^{\infty} \frac{du}{A^2/4 + (2\pi\nu_0 u/c)^2} \cdot \exp\left\{-\left(\frac{\Delta\nu \cdot c}{\nu}\right)^2 \frac{m}{2k\varphi}\right\}$$

$$= (c/A\nu_0) \exp\left\{-\left(\frac{\Delta\nu \cdot c}{\nu}\right)^2 \frac{m}{2k\varphi}\right\}.$$

For all values of  $\nu - \nu_0$  the integral may be expressed as

$$\frac{c}{\pi A\nu_0} \int_{-\infty}^{\infty} \frac{\exp\{-(x-y)/\eta\}^2 \cdot dy}{1+y^2},$$

where  $x = 4\pi(\nu - \nu_0)/A$

$$\eta = (4\pi\nu_0/Ac)(2k\varphi/m)^{\frac{1}{2}}.$$

Values of this integral have been tabulated as functions of  $x$  and  $\eta$  by Born(5), although his table does not cover the complete range required for astrophysical purposes. A very useful approximation has been given by Rosseland(6). When  $\Delta\nu$  is not too large,

$$\int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) \cdot du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2} \approx (c/A\nu_0) \left\{ \exp(-\xi)^2 + \frac{4D}{\pi\xi^2} \int_0^{\xi} \zeta^2 \exp(-\zeta^2) \cdot d\zeta \right\},$$

where  $D = (Ac/4\pi\nu_0)(m/2k\varphi)^{\frac{1}{2}}$

$$\xi = \{(\nu - \nu_0)c/\nu_0\}(m/2k\varphi)^{\frac{1}{2}}.$$

The value of  $A$  for the H $\alpha$ , L $\alpha$ , and L $\beta$  lines is not greatly affected by collisions. For transitions involving the 3P state it is approximately  $1.9 \times 10^8$  per second, while for all the other transitions it is approximately  $6.5 \times 10^8$  per second.

Substituting numerical values in the right-hand side of the above expression, we get

$$c/A\nu_0\{\exp(-5.44 \times 10^{12}[\Delta\lambda/\lambda]^2\varphi^{-1}) + 1.45 \times 10^{-13}A\lambda(\lambda/\Delta\lambda)^2\varphi^{\frac{1}{2}} \int_0^{\xi} \zeta^2 \exp(-\zeta^2) d\zeta\}.$$

If  $\varphi = 10^4$  °K., we find that this expression may be replaced by  $(c/A\nu_0)\exp(-\xi^2)$  without introducing an error greater than 10 per cent. for  $L\alpha$  and  $L\beta$  provided  $\Delta\lambda/\lambda < 1.2 \times 10^{-4}$ , while for  $H\alpha$  the corresponding values of  $\Delta\lambda/\lambda$  are  $1.0 \times 10^{-4}$  for  $3S \rightleftharpoons 2P$  and  $3D \rightleftharpoons 2P$  transitions, and  $1.1 \times 10^{-4}$  for  $3P \rightleftharpoons 2S$  transitions. These values of  $\Delta\lambda/\lambda$  increase with  $\varphi$ . Hence, provided  $\Delta\lambda$  lies within these ranges, we find that

$$\frac{d\Omega}{\pi} I_\nu d\nu = \frac{d\Omega}{4\pi} \left( \frac{m}{2\pi k\varphi} \right)^{\frac{1}{2}} ch.d\nu \{1 + \kappa / [\exp(-E_l^j/kT) - 1]\} \\ \times \exp\{ -[(\nu - \nu_0)c/\nu_0]^2 m/2k\varphi \} \sum N_j A_{j,l}.$$

### III. THE ABSORPTION COEFFICIENT

From the general equation for the radiation of energy we can deduce, by taking the absorption into account, the total energy radiated from the medium.

The atomic absorption coefficient  $\alpha_\nu$  is defined by

$$\tau = \exp[-(\alpha_\nu)_{l,j} N_l x],$$

where  $\tau$  is the transmission of a column of gas of length  $x$ , containing  $N_l$  atoms per unit volume. For stationary atoms,  $(\alpha_\nu)_{l,j}$  is given by

$$(\alpha_\nu)_{l,j} = \frac{c^2}{32\pi^3 \nu^2} \cdot \frac{\tilde{\omega}_j}{\tilde{\omega}_l} \cdot \frac{A_{j,l} A}{(\nu - \nu_a)^2 + A^2 / 16\pi^2},$$

where  $\tilde{\omega}_j$  and  $\tilde{\omega}_l$  are the respective statistical weights of the  $j$  and  $l$  states,\* and  $\nu_a$  is the frequency of the absorption maximum. This frequency differs slightly from  $\nu_0$ , but we make little error in following common practice and putting  $\nu_a = \nu_0$ .

For a Maxwellian distribution of atom velocities at a temperature  $\varphi$ , it is well known that  $(\alpha_\nu)_{l,j}$  is given by

$$(\alpha_\nu)_{l,j} = \frac{c^2 \tilde{\omega}_j A_{j,l} A}{8\pi \nu^2 \tilde{\omega}_l (2\pi k\varphi/m)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2}.$$

For sufficiently small values of  $\nu - \nu_0$ ,

$$(\alpha_\nu)_{l,j} = \frac{c^3 \tilde{\omega}_j A_{j,l}}{8\pi \nu^3 \tilde{\omega}_l (2\pi k\varphi/m)^{\frac{1}{2}}} \frac{\exp\{-(\nu - \nu_0)^2 mc^2/2k\varphi\nu_0^2\}}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2}.$$

In general, there may be several substates  $l$  from which an atom may absorb radiation of frequency  $\nu$ , so that the transmission is given by

$$\tau = \exp[-x \cdot \sum_l (\alpha_\nu)_{l,j} N_l].$$

\* MENZEL, D. H.—Lick Obs. Publ. **17**: 227 (1931), where this expression due to Jacqueline R. Oppenheimer is explicitly given.

## IV. THE PROFILE OF THE SPECTRAL LINE

If at a distance between  $x$  and  $x+dx$  from an observer, energy  $d\Omega \cdot I_\nu dx d\nu/4\pi$  is emitted per unit projected area into a solid angle  $d\Omega$ , the amount remaining after traversing the distance  $x$  is

$$\frac{d\Omega dx d\nu}{4\pi} \cdot I_\nu \exp\left[-\int_0^x \sum_l (\alpha_\nu)_{l,j} N_l \cdot dx\right].$$

From well-known photometric principles, this is also the amount of such radiation received from a solid angle  $d\Omega$  per unit projected area near the observer. The energy in the frequency interval  $d\nu$  received per unit area per second from the emitting region is thus

$$R_\nu d\nu d\Omega = \frac{d\nu d\Omega}{4\pi} \int_0^\infty I_\nu \exp\left[-\int_0^x \sum_l (\alpha_\nu)_{l,j} N_l \cdot dx\right] dx.$$

This equation gives formally the profile of the spectral line, whether in absorption or emission.

(a) In the special case where  $I_\nu$ ,  $(\alpha_\nu)_{l,j}$ , and  $N_l$  are constant for  $x < x_1$ , and zero for  $x > x_1$ , we have

$$R_\nu d\nu d\Omega = \frac{d\nu d\Omega I_\nu \{1 - \exp[-\sum_l (\alpha_\nu)_{l,j} N_l x_1]\}}{4\pi \sum_l (\alpha_\nu)_{l,j} N_l}.$$

Further

$$\frac{I_\nu}{\sum_l (\alpha_\nu)_{l,j} N_l} = \sum_j \left\{ \frac{(I_\nu)_{j,l}}{\sum_l (\alpha_\nu)_{l,j} N_l} \right\},$$

the right-hand side of which, on expanding and cancelling, becomes

$$\frac{E_l^j \{1 + \kappa / [\exp(E_l^j/kT) - 1]\} \sum_j A_{j,l} A \cdot N_j \int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2}}{\frac{c^2}{8\pi\nu^2} \sum_l \frac{\tilde{\omega}_j A_{j,l} A \cdot N_l}{\tilde{\omega}_l} \int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2}}.$$

(b) Now suppose that in addition the populations of the states  $j$  and  $l$  are distributed as in thermodynamic equilibrium at a temperature  $\zeta$ , so that

$$\tilde{\omega}_l N_j / \tilde{\omega}_j N_l = \exp(-E_l^j/k\zeta);$$

then, on substituting in the above expressions,

$$R_\nu d\nu d\Omega = \frac{2\nu^2 E_l^j d\nu d\Omega}{c^2 \exp(E_l^j/k\zeta)} [1 + \kappa / \{\exp(E_l^j/kT) - 1\}] [1 - \exp\{-\sum_l (\alpha_\nu)_{l,j} N_l x_1\}].$$



(c) If now  $\kappa \rightarrow 1$ ,  $\zeta \rightarrow T$ , so that the gas is effectively in thermodynamic equilibrium, we obtain the familiar result

$$R_{\nu} d\nu d\Omega = \frac{2\nu^2 E_l^j d\nu d\Omega [1 - \exp\{-\sum_l (\alpha_{\nu})_{l,j} N_l x_1\}]}{c^2 \{\exp(E_l^j/kT) - 1\}} \\ = P_{\nu,T} [1 - \exp\{-\sum_l (\alpha_{\nu})_{l,j} N_l x_1\}] d\nu d\Omega,$$

where  $P_{\nu,T}$  represents the radiation from a black body at temperature  $T$ .

(d) If the populations of the various states  $j$  and  $l$  are not distributed as in thermodynamic equilibrium at any temperature, we can still obtain a simple expression for  $R_{\nu} d\nu d\Omega$  under some conditions; for, when  $\nu - \nu_0$  is small, and lies within limits corresponding to those given earlier for  $\Delta\lambda$ ,

$$\int_{-\infty}^{\infty} \frac{\exp(-mu^2/2k\varphi) \cdot du}{A^2/4 + 4\pi^2(\nu - \nu_0 + \nu_0 u/c)^2} = \frac{c}{A\nu_0} \cdot \exp\{-[(\nu - \nu_0)c/\nu_0]^2 m/2k\varphi\}.$$

Then

$$R_{\nu} d\nu d\Omega = \frac{d\nu d\Omega (1 - \exp[-\sum_l (\alpha_{\nu})_{l,j} N_l x_1]) E_l^j \{1 + \kappa / [\exp(E_l^j/kT) - 1]\}}{c^2/2\nu^2} \cdot \frac{\sum_j A_{j,l} N_j}{\sum_l (\tilde{\omega}_j A_{j,l} N_l / \tilde{\omega}_l)}$$

For reasons which appear below, we now define the effective temperature of the gas  $\zeta$  by the relation,

$$\sum_j A_{j,l} N_j / \sum_l (\tilde{\omega}_j A_{j,l} N_l / \tilde{\omega}_l) = \exp(-E_l^j/k\zeta),$$

whence we obtain the result

$$R_{\nu} d\nu d\Omega = \frac{2\nu^2 E_l^j d\nu d\Omega}{c^2 \exp(E_l^j/k\zeta)} [1 + \kappa / \{\exp(E_l^j/kT) - 1\}] [1 - \exp\{-\sum_l (\alpha_{\nu})_{l,j} N_l x_1\}].$$

(e) If  $\exp(E_l^j/kT) \gg 1$ , so that  $1/\{\exp(E_l^j/kT) - 1\} \ll 1$ , then the result in (d) reduces to

$$R_{\nu} d\nu d\Omega = \frac{2\nu^2 E_l^j d\nu d\Omega [1 - \exp\{-\sum_l (\alpha_{\nu})_{l,j} N_l x_1\}]}{c^2 \exp(E_l^j/k\zeta)},$$

and, if  $\exp(E_l^j/k\zeta) \gg 1$ , this simplifies to

$$R_{\nu} d\nu d\Omega = P_{\nu,\zeta} [1 - \exp\{-\sum_l (\alpha_{\nu})_{l,j} N_l x_1\}] d\nu d\Omega.$$

This equation, which is valid only over the range of  $\nu - \nu_0$  indicated in (d), may be compared with the well-known result given in (c). We see immediately that the equations are identical in form, thus justifying our definition of effective temperature.

(f) Finally, we extend the above results to the case of a non-homogeneous medium. Every point of this medium will be supposed to be characterized by an effective temperature  $\zeta$ , and for simplicity we shall write  $\alpha_{\nu,x}$  to represent the value of  $\sum_l (\alpha_{\nu})_{l,j} N_l$  at  $x$ . Provided  $\nu - \nu_0$  is restricted so that the condition of (d) is everywhere satisfied, we find that the intensity of radiation originating in a layer  $dx$  is, on emerging from the layer, given by

$$P_{\nu,\zeta} [1 - \exp(-\alpha_{\nu,x} dx)] = P_{\nu,\zeta} \alpha_{\nu,x} dx.$$

At the plane  $x=x_1$ , absorption reduces this radiation by a factor  $\exp(-\int_{x_1}^x \alpha_{\nu,x} dx)$ ; integrating throughout the gas, we obtain

$$R_{\nu} d\nu d\Omega = d\nu d\Omega \int_{x_1}^{\infty} P_{\nu,\zeta} \alpha_{\nu,x} \exp(-\int_0^x \alpha_{\nu,x} dx) \cdot dx.$$

This equation is fundamentally the same as that given at the beginning of this section.

### V. TABLES OF EFFECTIVE TEMPERATURES

From the above expressions and with the data of papers I and II it is now possible to compute the intensities and profiles of some of the spectral lines of hydrogen for known physical conditions. In general the radiation, electron, and

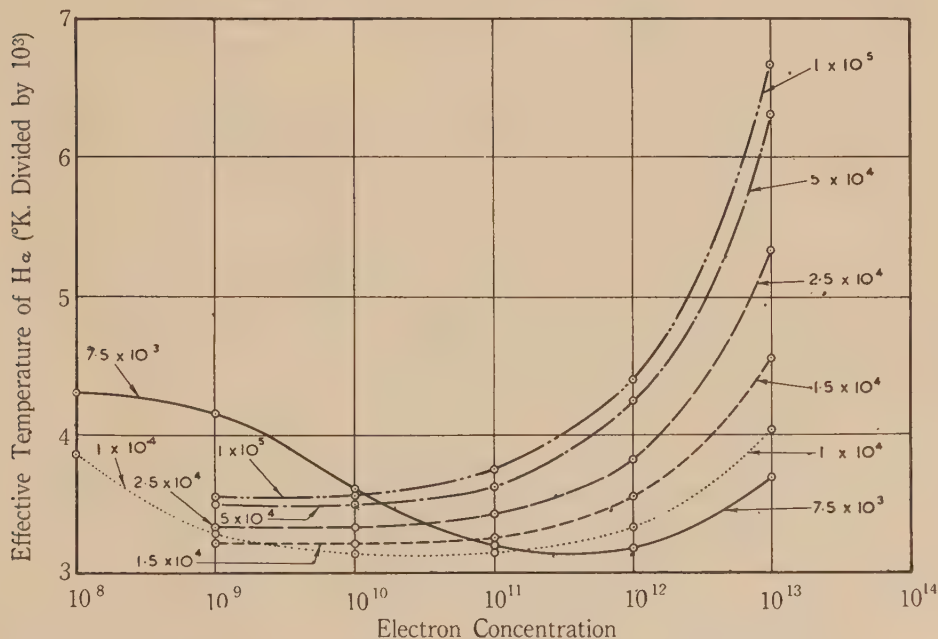


Fig. 1.—The effective temperature of H $\alpha$  plotted against electron concentration for a range of gas temperature from  $7.5 \times 10^3$  °K. to  $1.0 \times 10^5$  °K. Irradiation conditions:  $T=5 \times 10^3$  °K.,  $\kappa=0.5$ .

atom temperatures will all be different. In the absence of a sufficiently strong electric field, however, the gas will be in local thermal equilibrium so that both the electrons and the atoms will be at the same temperature, which may be called gas temperature. The results tabulated below apply to such a case.

The spectral emission is fully defined by the effective gas temperature and the absorption coefficient. Tables 1 to 6 give the effective temperature  $\zeta$  for

the  $L\alpha$ ,  $L\beta$ , and  $H\alpha$  lines, as defined in (d) and computed from the data contained in paper II. These data consist of the populations of the various substates of  $n=1, 2$ , and 3 of hydrogen for various electron temperatures in the range  $7.5 \times 10^3$  °K. to  $1.0 \times 10^5$  °K., and in the presence of radiation from a black body at a temperature  $T=5 \times 10^3$  °K. diluted by a factor  $\kappa$ . The concentrations of the atoms as given in paper II were considered to be uncertain to within a factor of the order of 2. If  $\sum_j A_{j,l} N_j / \sum_l (\tilde{\omega}_j A_{j,l} N_l / \tilde{\omega}_l)$  is too large or too small by a factor of 2, the estimated effective temperature is respectively too large or too small by an amount depending on the effective temperature, but of the order of 15 per cent. for  $\zeta \approx 4 \times 10^3$  °K.

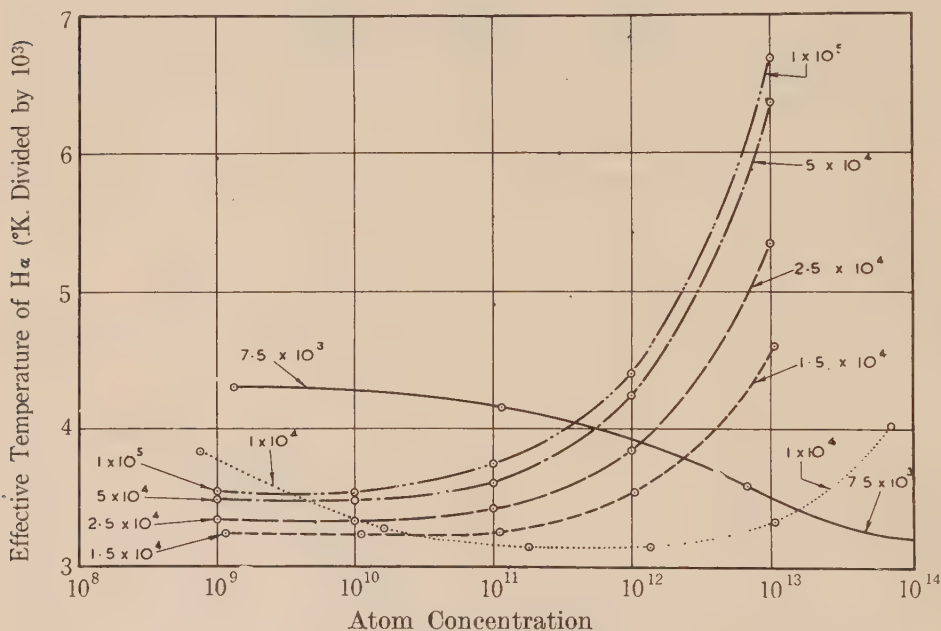


Fig. 2.—The effective temperature of  $H\alpha$  plotted against atom concentration (neutral atoms+ions) for a range of gas temperature from  $7.5 \times 10^3$  °K. to  $1.0 \times 10^5$  °K.

Irradiation conditions:  $T=5 \times 10^3$  °K.,  $\kappa=0.5$ .

The effective temperatures of  $H\alpha$  are plotted, for  $\kappa=0.5$ , firstly against electron concentration (Fig. 1), then against atom concentration (atom concentration=concentration of neutral atoms+concentration of ions) (Fig. 2), and finally against pressure (Fig. 3). In this last case, the unit of pressure is particles (electrons+ions+neutral atoms) per cc.  $\times$  °K.

From an inspection of the results, it will be noted that the effective temperature for  $L\beta$  is always greater than that for  $L\alpha$ , and both considerably exceed the effective temperature for  $H\alpha$ . The latter shows an interesting variation with the gas temperature. As the gas temperature is raised, the effective temperature of  $H\alpha$  first decreases, owing to the relatively greater

concentration of atoms in the  $2S$  state. The effective temperature is a minimum when the gas temperature is about  $1.0 \times 10^4$  to  $1.5 \times 10^4$  °K., and then increases with increasing gas temperature.

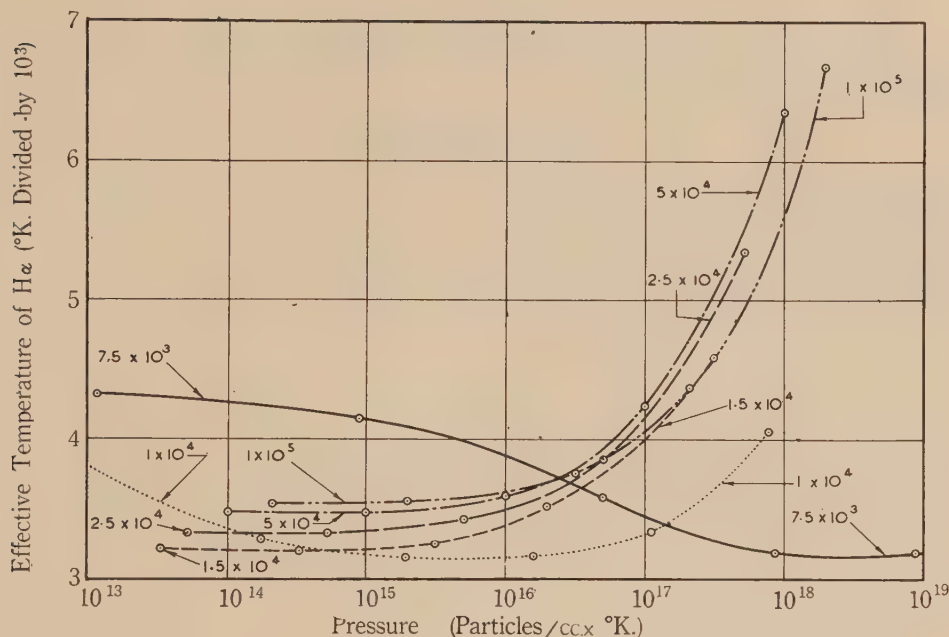


Fig. 3.—The effective temperature of  $H\alpha$  plotted against pressure, measured as the total number of particles (neutral atoms + ions + electrons) per cc.  $\times$  °K., for a range of gas temperature from  $7.5 \times 10^3$  °K. to  $1.0 \times 10^5$  °K.

Irradiation conditions:  $T = 5 \times 10^3$  °K.,  $\alpha = 0.5$ .

## VI. TABLES OF ABSORPTION COEFFICIENTS

We may note that the absorption is determined by the absorption coefficient  $\sum_l [(\alpha_v)_{lj} N_l]$ . When  $v - v_0$  is small, this expression equals

$$\frac{c^3 \exp\{-(v - v_0)^2 c^2 m / 2k\varphi v_0^2\}}{8\pi v^3 (2\pi k\varphi/m)^{\frac{1}{2}}} \sum_l \frac{\tilde{\omega}_j A_{j,l} N_l}{\tilde{\omega}_l}$$

At the centre of the absorption line  $v = v_0$ , and the expression simplifies to

$$\frac{1}{8\pi(2\pi k/m)^{\frac{1}{2}}} \cdot \frac{\lambda^3}{\varphi^{\frac{1}{2}}} \sum_l \frac{\tilde{\omega}_j A_{j,l} N_l}{\tilde{\omega}_l}$$

On substituting numerical values for the physical constants, this becomes

$$1.75 \times 10^{-6} (\lambda^3 / \varphi^{\frac{1}{2}}) \sum_l (\tilde{\omega}_j A_{j,l} N_l / \tilde{\omega}_l)$$

Values of the absorption coefficients at the centre of the  $L\alpha$  and  $H\alpha$  lines, as computed from this formula with the data of paper II, together with the absorption coefficients of hydrogen when in thermodynamic equilibrium at  $5 \times 10^3$  °K., are given in Tables 7 and 8 as functions of the electron concentration. The absorption coefficient at the centre of  $L\beta$  is 0.267 of that at the centre of  $L\alpha$

and may thus be obtained readily from Table 7. Since the concentrations of the atoms in the different states are uncertain to within a factor of the order of 2, the absorption coefficients are uncertain to the same extent.

### VII. ACKNOWLEDGMENT

This work has been carried out as part of the programme of the Division of Physics, C.S.I.R.

### VIII. REFERENCES

- (1) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1**: 275 (1948).
- (2) GIOVANELLI, R. G.—*Ibid.* **1**: 289 (1948).
- (3) WEISSKOPF, V., and WIGNER, E.—*Z. Phys.* **63**: 54 (1930).
- (4) WEISSKOPF, V.—*Phys. Z.* **34**: 1 (1933).
- (5) BORN, M.—“*Optik*”, p. 486. (Springer: Berlin, 1933.)
- (6) ROSSELAND, S.—“*Theoretical Astrophysics*”, p. 101. (Clarendon Press: Oxford, 1936.)

TABLE 1

EFFECTIVE TEMPERATURES OF HYDROGEN

Radiation Temperature  $T=5 \times 10^3$  °K. Electron Concentration  $N_e=10^{18}$  per cc.

Gas Temperature (° K.)	Effective Temperature (divided by $10^3$ )								
	H $\alpha$			L $\alpha$			L $\beta$		
	$\kappa=1$	0.5	0	1	0.5	0	1	0.5	0
$7.5 \times 10^3$	3.88	3.70	3.46	5.23	5.20	5.18	6.14	6.03	5.67
$1.0 \times 10^4$	4.18	4.04	3.89	6.20	6.20	6.20	7.40	7.28	6.90
$1.5 \times 10^4$	4.70	4.57	4.45	7.75	7.75	7.75	9.24	9.08	8.74
$2.5 \times 10^4$	5.42	5.37	5.25	9.64	9.67	9.68	11.3	11.1	10.8
$5.0 \times 10^4$	6.41	6.34	6.29	11.8	11.8	11.8	13.6	13.4	13.2
$1.0 \times 10^5$	6.72	6.69	6.58	13.2	13.2	13.2	15.3	15.2	14.9



TABLE 2

 $T=5 \times 10^3$  °K.  $N_e=10^{12}$  per cc.

Gas Temperature (°K.)	Effective Temperature (divided by $10^3$ )								
	H $\alpha$			L $\alpha$			L $\beta$		
	$\kappa=1$	0.5	0	1	0.5	0	1	0.5	0
$7.5 \times 10^3$	3.56	3.21	2.65	5.03	4.93	4.70	5.80	5.78	5.16
$1.0 \times 10^4$	3.60	3.33	2.87	5.58	5.53	5.53	6.88	6.80	6.19
$1.5 \times 10^4$	3.78	3.52	3.12	6.70	6.70	6.76	8.42	8.33	7.60
$2.5 \times 10^4$	4.05	3.82	3.51	8.09	8.11	8.18	10.1	9.98	9.12
$5.0 \times 10^4$	4.50	4.25	3.99	9.55	9.60	9.62	11.9	11.7	10.8
$1.0 \times 10^5$	4.64	4.40	4.10	10.4	10.5	10.6	13.1	12.9	11.9

TABLE 3

 $T=5 \times 10^3$  °K.  $N_e=10^{11}$  per cc.

Gas Temperature (°K.)	Effective Temperature (divided by $10^3$ )								
	H $\alpha$			L $\alpha$			L $\beta$		
	$\kappa=1$	0.5	0	1	0.5	0	1	0.5	0
$7.5 \times 10^3$	3.76	3.21	2.10	5.00	4.87	4.30	5.36	5.33	4.77
$1.0 \times 10^4$	3.52	3.17	2.23	5.13	5.08	5.00	6.23	6.21	5.59
$1.5 \times 10^4$	3.60	3.26	2.38	5.92	5.92	5.99	7.49	7.48	6.76
$2.5 \times 10^4$	3.80	3.42	2.59	6.97	6.98	7.07	8.86	8.82	7.92
$5.0 \times 10^4$	4.00	3.61	2.85	8.00	8.00	8.14	10.2	10.2	9.19
$1.0 \times 10^5$	4.10	3.75	2.92	8.62	8.62	8.77	11.2	11.0	9.98

TABLE 4

 $T=5 \times 10^3$  °K.  $N_e=10^{10}$  per cc.

Gas Temperature (°K.)	Effective Temperature (divided by $10^3$ )								
	H $\alpha$			L $\alpha$			L $\beta$		
	$\kappa=1$	0.5	0	1	0.5	0	1	0.5	0
$7.5 \times 10^3$	4.50	3.61	1.72	5.00	4.87	4.12	5.09	5.01	4.42
$1.0 \times 10^4$	3.58	3.15	1.81	5.01	4.90	4.56	5.68	5.67	5.11
$1.5 \times 10^4$	3.60	3.23	1.91	5.36	5.33	5.34	6.68	6.68	6.06
$2.5 \times 10^4$	3.75	3.35	2.05	6.13	6.13	6.21	7.85	7.85	7.00
$5.0 \times 10^4$	3.92	3.51	2.20	6.91	6.91	7.02	8.82	8.82	8.00
$1.0 \times 10^5$	4.02	3.57	2.25	7.38	7.38	7.51	9.50	9.50	8.58

TABLE 5

 $T=5 \times 10^3$  °K.  $N_e=10^9$  per cc.

Gas Temperature (°K.)	Effective Temperature (divided by $10^3$ )								
	H $\alpha$			L $\alpha$			L $\beta$		
	$\kappa=1$	0.5	0	1	0.5	0	1	0.5	0
$7.5 \times 10^3$	4.94	4.17	1.46	5.00	4.87	3.69	5.01	4.89	4.12
$1.0 \times 10^4$	4.07	3.30	1.52	5.00	4.88	4.19	5.24	5.21	4.72
$1.5 \times 10^4$	3.62	3.23	1.59	5.10	4.99	4.86	6.02	6.01	5.52
$2.5 \times 10^4$	3.75	3.35	1.68	5.50	5.49	5.55	6.90	6.90	6.30
$5.0 \times 10^4$	3.91	3.50	1.78	6.10	6.10	6.20	7.71	7.71	7.06
$1.0 \times 10^5$	4.03	3.56	1.92	6.48	6.47	6.57	8.20	8.20	7.50

TABLE 6

 $T=5 \times 10^3$  °K.  $N_e=10^8$  per cc.

Gas Temperature (°K.)	Effective Temperature (divided by $10^3$ )								
	H $\alpha$			L $\alpha$			L $\beta$		
	$\kappa=1$	0.5	0	1	0.5	0	1	0.5	0
$7.5 \times 10^3$	5.01	4.31	1.26	5.00	4.87	3.41	5.00	4.88	3.87
$1.0 \times 10^4$	4.73	3.85	1.32	5.00	4.88	3.87	5.05	4.96	4.41

TABLE 7

ABSORPTION COEFFICIENTS AT THE CENTRE OF L $\alpha$ Radiation Temperature  $T'=5 \times 10^3$  °K.

Electron Conc. $N_e$ (per cc.)	Dilution Factor $\kappa$	Gas Temperature (°K.)						
		$5.0 \times 10^3$	$7.5 \times 10^3$	$1.0 \times 10^4$	$1.5 \times 10^4$	$2.5 \times 10^4$	$5.0 \times 10^4$	$1.0 \times 10^5$
$10^{13}$	1	$5.4 \times 10^5$	$6.4 \times 10^2$	3.1	$1.3 \times 10^{-2}$	$1.6 \times 10^{-4}$	$3.2 \times 10^{-6}$	$2.8 \times 10^{-7}$
	0.5		$6.6 \times 10^2$	2.9	$1.2 \times 10^{-2}$	$1.5 \times 10^{-4}$	$3.2 \times 10^{-6}$	$2.8 \times 10^{-7}$
	0		$6.9 \times 10^2$	2.7	$1.1 \times 10^{-2}$	$1.4 \times 10^{-4}$	$3.1 \times 10^{-6}$	$2.7 \times 10^{-7}$
$10^{12}$	1	$5.4 \times 10^3$	$7.9 \times 10$	$6.0 \times 10^{-1}$	$2.8 \times 10^{-3}$	$2.8 \times 10^{-5}$	$4.4 \times 10^{-7}$	$3.4 \times 10^{-8}$
	0.5		$7.6 \times 10$	$5.1 \times 10^{-1}$	$2.2 \times 10^{-3}$	$2.2 \times 10^{-5}$	$3.9 \times 10^{-7}$	$3.2 \times 10^{-8}$
	0		$6.8 \times 10$	$2.6 \times 10^{-1}$	$1.1 \times 10^{-3}$	$1.5 \times 10^{-5}$	$3.2 \times 10^{-7}$	$2.8 \times 10^{-8}$
$10^{11}$	1	$5.4 \times 10$	6.8	$9.3 \times 10^{-2}$	$5.6 \times 10^{-4}$	$4.9 \times 10^{-6}$	$6.4 \times 10^{-8}$	$4.0 \times 10^{-9}$
	0.5		8.0	$9.2 \times 10^{-2}$	$5.2 \times 10^{-4}$	$4.4 \times 10^{-6}$	$6.0 \times 10^{-8}$	$4.0 \times 10^{-9}$
	0		7.2	$2.7 \times 10^{-2}$	$1.1 \times 10^{-4}$	$1.5 \times 10^{-6}$	$3.2 \times 10^{-8}$	$2.8 \times 10^{-9}$
$10^{10}$	1	$5.4 \times 10^{-1}$	$2.5 \times 10^{-1}$	$9.6 \times 10^{-3}$	$6.3 \times 10^{-5}$	$5.5 \times 10^{-7}$	$6.9 \times 10^{-9}$	$4.2 \times 10^{-10}$
	0.5		$4.4 \times 10^{-1}$	$1.0 \times 10^{-2}$	$6.4 \times 10^{-5}$	$5.5 \times 10^{-7}$	$6.9 \times 10^{-9}$	$4.2 \times 10^{-10}$
	0		$7.2 \times 10^{-1}$	$2.8 \times 10^{-3}$	$1.1 \times 10^{-5}$	$1.5 \times 10^{-7}$	$3.2 \times 10^{-9}$	$2.8 \times 10^{-10}$
$10^9$	1	$5.4 \times 10^{-3}$	$3.4 \times 10^{-3}$	$7.2 \times 10^{-4}$	$6.5 \times 10^{-6}$	$5.6 \times 10^{-8}$	$7.0 \times 10^{-10}$	$4.2 \times 10^{-11}$
	0.5		$8.0 \times 10^{-3}$	$9.0 \times 10^{-4}$	$6.6 \times 10^{-6}$	$5.6 \times 10^{-8}$	$7.1 \times 10^{-10}$	$4.3 \times 10^{-11}$
	0		$7.2 \times 10^{-2}$	$2.8 \times 10^{-4}$	$1.1 \times 10^{-6}$	$1.5 \times 10^{-8}$	$3.2 \times 10^{-10}$	$2.8 \times 10^{-11}$
$10^8$	1	$5.4 \times 10^{-5}$	$3.5 \times 10^{-5}$	$2.0 \times 10^{-5}$				
	0.5		$8.8 \times 10^{-5}$	$3.9 \times 10^{-5}$				
	0		$7.2 \times 10^{-3}$	$2.8 \times 10^{-5}$				

TABLE 8  
 ABSORPTION COEFFICIENTS AT THE CENTRE OF H $\alpha$   
 Radiation Temperature  $T=5 \times 10^3$  °K.

Electron Conc. $N_e$ (per cc.)	Dilu- tion Factor $\kappa$	Gas Temperature (°K.)						
		$5.0 \times 10^3$	$7.5 \times 10^3$	$1.0 \times 10^4$	$1.5 \times 10^4$	$2.5 \times 10^4$	$5.0 \times 10^4$	$1.0 \times 10^5$
$10^{13}$	1	$9.2 \times 10^{-4}$	$3.1 \times 10^{-4}$	$6.3 \times 10^{-5}$	$9.5 \times 10^{-6}$	$1.3 \times 10^{-6}$	$1.6 \times 10^{-7}$	$3.6 \times 10^{-8}$
	0.5		$3.8 \times 10^{-4}$	$7.1 \times 10^{-5}$	$1.0 \times 10^{-5}$	$1.3 \times 10^{-6}$	$1.7 \times 10^{-7}$	$3.7 \times 10^{-8}$
	0		$4.8 \times 10^{-4}$	$8.0 \times 10^{-5}$	$1.1 \times 10^{-5}$	$1.4 \times 10^{-6}$	$1.7 \times 10^{-7}$	$3.9 \times 10^{-8}$
$10^{12}$	1	$9.2 \times 10^{-6}$	$9.0 \times 10^{-6}$	$3.2 \times 10^{-6}$	$6.0 \times 10^{-7}$	$9.3 \times 10^{-8}$	$1.1 \times 10^{-8}$	$2.2 \times 10^{-9}$
	0.5		$1.6 \times 10^{-5}$	$4.6 \times 10^{-6}$	$7.6 \times 10^{-7}$	$1.1 \times 10^{-7}$	$1.3 \times 10^{-8}$	$2.8 \times 10^{-9}$
	0		$4.7 \times 10^{-5}$	$7.7 \times 10^{-6}$	$1.0 \times 10^{-6}$	$1.4 \times 10^{-7}$	$1.7 \times 10^{-8}$	$3.8 \times 10^{-9}$
$10^{11}$	1	$9.2 \times 10^{-8}$	$1.1 \times 10^{-7}$	$5.8 \times 10^{-8}$	$1.5 \times 10^{-8}$	$2.3 \times 10^{-9}$	$2.7 \times 10^{-10}$	$4.8 \times 10^{-11}$
	0.5		$2.3 \times 10^{-7}$	$1.2 \times 10^{-7}$	$2.6 \times 10^{-8}$	$4.0 \times 10^{-9}$	$4.6 \times 10^{-10}$	$7.8 \times 10^{-11}$
	0		$5.0 \times 10^{-6}$	$8.1 \times 10^{-7}$	$1.1 \times 10^{-7}$	$1.4 \times 10^{-8}$	$1.7 \times 10^{-9}$	$3.8 \times 10^{-10}$
$10^{10}$	1	$9.2 \times 10^{-10}$	$7.9 \times 10^{-10}$	$6.3 \times 10^{-10}$	$1.7 \times 10^{-10}$	$2.8 \times 10^{-11}$	$3.1 \times 10^{-12}$	$5.4 \times 10^{-13}$
	0.5		$1.7 \times 10^{-9}$	$1.4 \times 10^{-9}$	$3.4 \times 10^{-10}$	$5.4 \times 10^{-11}$	$6.2 \times 10^{-12}$	$1.1 \times 10^{-12}$
	0		$5.0 \times 10^{-7}$	$8.1 \times 10^{-8}$	$1.1 \times 10^{-8}$	$1.4 \times 10^{-9}$	$1.7 \times 10^{-10}$	$3.8 \times 10^{-11}$
$10^9$	1	$9.2 \times 10^{-12}$	$6.4 \times 10^{-12}$	$5.9 \times 10^{-12}$	$1.8 \times 10^{-12}$	$2.8 \times 10^{-13}$	$3.2 \times 10^{-14}$	$5.5 \times 10^{-15}$
	0.5		$1.0 \times 10^{-11}$	$1.2 \times 10^{-11}$	$3.6 \times 10^{-12}$	$5.6 \times 10^{-13}$	$6.4 \times 10^{-14}$	$1.1 \times 10^{-14}$
	0		$5.0 \times 10^{-8}$	$8.2 \times 10^{-9}$	$1.1 \times 10^{-9}$	$1.4 \times 10^{-10}$	$1.7 \times 10^{-11}$	$3.8 \times 10^{-12}$
$10^8$	1	$9.2 \times 10^{-14}$	$6.1 \times 10^{-14}$	$4.8 \times 10^{-14}$				
	0.5		$9.2 \times 10^{-14}$	$9.0 \times 10^{-14}$				
	0		$5.0 \times 10^{-9}$	$8.2 \times 10^{-10}$				

# STATISTICAL MECHANICS OF HIGH POLYMER SOLUTIONS

By A. R. MILLER\*

[Manuscript received August 16, 1948]

## Summary

The configurational partition function for random mixtures containing any numbers of components which can consist of simple, simple chain, branched chain, or closed ring molecules is examined. Using a general statistical method a set of partial differential equations is obtained for the appropriate combinatorial factor. The integrability of this set of equations is examined. This provides a criterion by which it can be decided in what cases a mathematically precise value for the combinatorial factor can be obtained rigorously by solving the set of partial differential equations. In these cases general formulae are given for the combinatorial factor. It is shown that precise formulae can be deduced rigorously (a) for all binary mixtures whether the components consist of simple, simple chain, branched chain, or closed ring molecules; (b) mixtures of any number of components containing not more than one high polymer species, which can consist of closed ring equally well as of simple chain or branched chain molecules; and (c) mixtures of any number of components containing more than one high polymer species provided these consist only of simple chain or branched chain molecules.

It is also shown that even in the case in which the condition of integrability is not satisfied, an approximation, which involves negligible error for high polymer molecules, indicates that the general formula must still provide a good approximation practically. The approximations inherent in the physical model are also considered.

## I. INTRODUCTION

(a) It is a fact of common experience that in the presence of a solvent the behaviour of a high polymer, such as natural rubber, is markedly different from that of low molecular substances. When a quantitative experimental study of the properties of high polymer-solvent systems is carried out it is found that they show very great divergences from the mathematically simple laws which describe the behaviour of ideal solutions. The behaviour of such systems can, of course, be completely determined once the appropriate partition function has been constructed.

(b) This has been attempted by a direct enumeration of the distinct configurations of the assembly(1, 2, 3), following a method used by Fowler and Rushbrooke(4) to consider a mixture of monomer and dimer molecules. Such an enumeration, however, introduces unwanted algebraic complications and involves taking account of irrelevant factors (Miller, 5, pp. 62-3; Guggenheim, 6, pp. 208-10). An unambiguous formula for the partition function was first given(5, 7) using the approximation of local configurations which had been introduced in the theory of order-disorder in alloys(8). This involved a mathe-

\* Royal Society Mond Laboratory, Cavendish Laboratory, Cambridge.



mathematical extrapolation from the results for more tractable systems to systems containing polymer molecules. A general kinetic argument was subsequently given by Guggenheim(6) who stated, without proof, that his results were limited to certain kinds of molecules. A critical account and comparison of these methods has been given elsewhere(9).

In the present paper a general formula, applicable to mixtures containing any number of components, is derived on the basis of statistical arguments, and the validity of the results, depending upon the integrability of a set of partial differential equations, is examined specifically for different kinds of polymer molecules (simple chain, branched chain, and closed ring).

## II. LIQUID MODEL

(a) It is first necessary to describe the physical model in terms of which the analysis is carried out. The simplest useful model of a liquid of nearly spherical molecules is that of spheres packed together, not tightly which would represent a crystalline solid, but not too far from tightly. Such a model has considerably less order than a crystalline solid but it possesses a degree of order which makes it far different from the almost complete randomness of configuration which characterizes a gas. This model, which is called a quasi-crystalline array, ensures that the individual molecules are allowed considerable freedom of movement; at the same time it secures a high degree of local order in the geometrical relationship between a molecule and its near neighbours. At temperatures well below the critical temperature the number of neighbours surrounding each molecule in a liquid has a well-defined average value. There are fluctuations about this average value but they are not serious: except for an unimportant fraction of centres in the liquid, the regularity of the geometrical relationship between each molecule and its closest neighbours will not be disturbed by them. The fluctuations are, however, sufficient to destroy any long-range order, so that there is no regular geometrical relationship between a molecule and its more distant neighbours. This is the simplest physical model which is in agreement with the available X-ray data on liquids (for a review of this data see Gingrich,10, and for later work, Finbak,11).

(b) In considering mixtures which contain high polymer molecules this physical model can be adapted in the following way. The polymer molecule is formed by the repetition of a number of segments or submolecules. In the typical case of natural rubber, for instance, the segment is an isoprene unit, and it is of roughly the same size as a molecule of benzene or other typical solvent. We consider a quasi-crystalline array of sites as defined in the preceding paragraph. Each segment of the polymer molecule is allotted to one site of the array, and adjacent segments of the polymer molecule are allotted to pairs of closest neighbour sites of the array. Thus the molecule will occupy a small group of sites of the quasi-crystalline array, the sites of this group to be inter-related in such a way that proceeding from segment to segment of the molecule a path is traversed through the array of sites such that each site on the path is a closest neighbour of the immediately preceding and the immediately succeeding sites on the path. In this way simple chain, branched chain, and closed ring

molecules can be accommodated on the quasi-crystalline array and any degree of flexibility of the molecule about its bonds can be allowed for. Each solvent molecule can be considered to occupy one site of the array.

(c) In considering a physical model of this kind it is found that a quantity of importance is the number of closest neighbour contacts made by each polymer molecule, that is, the number of closest neighbour interactions into which each polymer molecule enters. Specifying this quantity by  $zq$ , where  $z$  is the average number of closest neighbour sites of any site on the array, its value can be calculated(12). For simple chain and branched chain molecules

$$(r-q)z=2(r-1) \quad \dots\dots\dots (1.1)$$

while for closed ring molecules

$$(r-q)z=2r \quad \dots\dots\dots (1.2)$$

where  $r$  is the number of segments in each polymer molecule. It must be remembered that  $r$  is very large, of the order of many thousands, so that in practice  $(r-1)/r$  differs negligibly from unity. Thus the values of  $q$  for the different cases differ insignificantly. The effect of this difference will be seen below.

(d) If in a mixture containing a number of components, quantities  $X_{ij}$  (all  $i, j$ ) are defined by the fact that  $zX_{ij}$  is the number of pairs of closest neighbour sites on the quasi-crystalline array of which one is occupied by a segment of a molecule  $i$  and the other by a segment of a molecule  $j$ , then

$$2X_{ii} + \sum'_k X_{ik} = q_i N_i \quad (\text{all } i) \quad \dots\dots\dots (2.1)$$

where  $N_i$  is the number of molecules  $i$  and the prime to the summation sign indicates that the summation is to be carried out over all  $k$  not equal to  $i$ . Again, if  $zx_{ii}$  is the number of pairs of closest neighbour sites which are occupied by adjacent segments of the same molecule  $i$  then

$$2x_{ii} = (r_i - q_i) N_i \quad (\text{all } i) \quad \dots\dots\dots (2.2)$$

These equations complete the description of the physical model and the specification of the parameters which describe the occupation of the sites of the quasi-crystalline array.

### III. PARTITION FUNCTION

(a) In considering a liquid mixture one is concerned with macrostates of an assembly which are differentiated by the configurational energy of the whole assembly; the factors in the partition function which correspond to the translational motion and to the internal degrees of freedom (rotational, vibrational, electronic, nuclear spin) of the individual systems of which the assembly is composed are taken as independent of the configuration of the assembly as a whole. Thus, one is not concerned to distinguish between states which, even though they may differ on a microscopical view, have the same configurational energy and so are identical macroscopically. For an assembly of  $N_i$  indistinguishable systems  $i$  (all  $i$ ) confined to a volume  $V$ , the partition function is given(13), with suitable choice of the energy zero, by a phase integral

$$Z = \prod_i \{l_i(T) j_i(T)\}^{N_i} (\prod_i N_i!)^{-1} \int \dots \int \exp\{-W(N_i, X_{ij}, x_{ii})/kT\} (d\omega_i)^{N_i} \dots (3)$$

Here  $l_i(T)$  is the partition function per molecule  $i$  corresponding to its translational motion,  $j_i(T)$  is that corresponding to the internal degrees of freedom,  $W(N_i, X_{ij}, x_{ii})$  is the configurational energy of the assembly,  $d\omega_i$  is an element of phase space corresponding to a molecule  $i$ , the integration is extended over the whole of phase space accessible to the individual molecules, the factorials are introduced to avoid multiple counting of indistinguishable configurations and the  $X_{ij}$  and the  $x_{ii}$  are the parameters which have been defined in Section II (*d*). Alternatively, in terms of the physical model which has been adopted, the partition function for the assembly can be written as

$$Z = g(N_i, X_{ij}, x_{ii}) \exp\{-W(N_i, X_{ij}, x_{ii})/kT\} \prod_i \varphi_i^{N_i} \dots \quad (4)$$

where  $\varphi_i$  is written for the factor in the partition function for a molecule  $i$  in the assembly corresponding to all its internal degrees of freedom and all forms of energy apart from the configurational energy and  $g(N_i, X_{ij}, x_{ii})$ , the combinatory factor, is the number of distinct configurations of molecules on the quasi-crystalline array of sites and having a configurational energy  $W(N_i, X_{ij}, x_{ii})$ . It has been shown(14) that the difficulty in the evaluation of a phase integral such as

$$\Omega(T) = \{\prod_i N_i!\}^{-1} \int \dots \int \exp\{-W(N_i, X_{ij}, x_{ii})/kT\} \prod_i (d\omega_i)^{N_i} \quad (5)$$

or of a combinatory factor

$$g(N_i, X_{ij}, x_{ii}) \dots \quad (6)$$

lies in the fact that in the specification of the occupation of the array of sites and of the corresponding configurational energy both the numbers of particles of different kinds, and the numbers of pairs of closest neighbour sites of different kinds (as determined by the occupation of the two sites of the pair) are involved. For approximately spherical molecules of the same size this difficulty vanishes if the energies of mixing are zero and then the mixture behaves as an ideal solution; but for mixtures containing polymer molecules the difficulty persists in all cases. This is so because even when the energies of mixing are zero, so that the  $X_{ij}$  become irrelevant to the specification of the assembly and its configurational energy, these quantities still involve the parameters  $x_{ii}$  which refer to pairs of sites.\*

(*b*) The case in which the energies of mixing vanish is, of course, still a simpler case with which to deal. Furthermore, it has the very great advantage that by consideration of this case it is possible to isolate the effects of size and shape alone. It has, in fact, been found (Miller, 9, Section 5.12) that these effects predominate and are responsible for almost the entire departure of solutions of high polymers from the laws of ideal solutions. Consequently, it also turns out that consideration of this case provides all the information of most interest and importance.

\* Ref. (9), closing paragraph of Ch. II.



## IV. STATISTICAL ANALYSIS

(a) It has been shown elsewhere (Miller, 12, Section 7) that for mixtures containing any number of components of any kind of molecule the probability of occupation of a given site  $A$  by a segment of a molecule  $i$  is

$$\frac{r_i N_i}{\sum_k r_k N_k} \dots \dots \dots (7)$$

and of a pair of closest neighbour sites  $A$  and  $B$ , one by a segment of a molecule  $i$  and the other by a segment of a molecule  $j$ , is

$$\frac{q_i q_j N_i N_j}{\sum_k r_k N_k \sum_k q_k N_k} \dots \dots \dots (8)$$

each being for a random distribution. Consider a mixture of a monomer and a polymer (containing  $r$  segments) in which the energy of mixing is zero so that there is a random distribution of molecules on the quasi-crystalline array. Consider a group of sites,  $\Gamma$ ,  $r$  in number, so interrelated that they can accommodate a polymer molecule. For a random arrangement of molecules the frequency of occupation of this group of sites by polymer molecules must be directly proportional to  $N_r$ ; it can be denoted by  $\kappa N_r / N$  where  $\kappa$  may depend on  $r$ ,  $q$ , and  $z$  but is independent of  $N_1$  and  $N_r$ . From equation (7), the probability that the end site of  $\Gamma$  is occupied by a monomer is

$$\frac{N_1}{N_1 + r N_r}$$

and from equation (8), the probability that this site and its closest neighbour which is also in  $\Gamma$  are both occupied by monomer molecules is

$$\frac{N_1}{N_1 + r N_r} \cdot \frac{N_1}{N_1 + q N_r}$$

Likewise, for each successive site in  $\Gamma$ , so that the probability that all the sites in  $\Gamma$  are occupied by monomer molecules is

$$\frac{N_1}{N_1 + r N_r} \left( \frac{N_1}{N_1 + q N_r} \right)^{r-1} \dots \dots \dots (9)$$

(b) Two states of the assembly of monomer and polymer molecules in contact with their vapour are now defined in the following way.

State  $\alpha$ :  $N_1$  monomer and  $N_r$  polymer molecules,

state  $\beta$ :  $N_1 + r$  monomer and  $N_r - 1$  polymer molecules.

For a random mixture the relative frequencies of occurrence of the states  $\alpha$  and  $\beta$ , denoted by  $P_\alpha$  and  $P_\beta$  respectively, are proportional simply to the number of distinct configurations of the given numbers of molecules of each kind. Thus

$$\frac{P_\alpha}{P_\beta} = \frac{g(N_1, N_r)}{g(N_1 + r, N_r - 1)} \dots \dots \dots (10)$$

an equation which merely makes use of the statistical weight. In the present case, the randomness of the mixture imposes a constraint and the sum over all the states of assembly reduces to the combinatory factor. In this case, the

combinatory factor is the degeneracy of the state which has the given (constant) value of the configurational energy, determined by the numbers of molecules of each species.

If the total number of sites on the array is kept fixed and  $N_r$  is decreased by one so that  $N_1$  is necessarily increased by  $r$  then the increment in  $\log g(N_1, N_r)$  is given by

$$\begin{aligned} [\Delta \log g(N_1, N_r)]_{\Delta N_r = -1, \Delta N_1 = +r} &= \log g(N_1 + r, N_r - 1) - \log g(N_1, N_r) \\ &= -\log \frac{g(N_1, N_r)}{(N_1 + r, N_r - 1)}. \end{aligned}$$

Thus, using equation (10), we obtain

$$\left[ \frac{\partial \log g(N_1, N_r)}{\partial N_r} \right]_N + \log \frac{P_\alpha}{P_\beta} = 0 \quad \dots \quad (11)$$

Now let  $N_r - 1$  polymer molecules and  $N_1$  monomer molecules be arranged in any configuration specified by  $C_r$ , on the sites outside  $\Gamma$  and let the probability of this configuration be  $P_r$ . The occupation of  $\Gamma$  by a polymer molecule is directly proportional to  $N_r$ . Thus, since there is complete randomness in the occupation of the sites, the relative frequency of the configuration in which a polymer molecule occupies  $\Gamma$  and the remaining  $N - r$  sites are occupied in the manner  $C_r$  is

$$P_r N_r / N \quad \dots \quad (12)$$

apart from a factor which may depend on  $r, q, z$ , and  $T$ , but is independent of  $N_1$  and  $N_r$ . This is simply the frequency of occurrence of the state  $\alpha$ . Again, from equation (9), the probability of the configuration in which monomer molecules occupy  $\Gamma$  and the remaining sites are occupied in the manner  $C_r$  is

$$P_r \frac{N_1}{N} \left( \frac{N_1}{N_1 + q N_r} \right)^{r-1} \quad \dots \quad (13)$$

This is simply the frequency of occurrence of the state  $\beta$ . Thus, the ratio of  $P_\alpha$  to  $P_\beta$  can be obtained from equations (12) and (13). The result is independent of  $P_r$ , that is, it is independent of the configuration of molecules on the sites outside  $\Gamma$  and so must be valid whatever that configuration. Thus

$$\rho \frac{P_\alpha}{P_\beta} = \frac{N_r}{N_1} \left( \frac{N_1 + q N_r}{N_1} \right)^{r-1} \quad \dots \quad (14)$$

where  $\rho$  may depend on  $r, q, z$ , and  $T$  but is independent of  $N_1$  and  $N_r$ . Equations (11) and (14) yield the differential equation

$$\begin{aligned} \left[ \frac{\partial}{\partial N_r} \log g(N_1, N_r) \right]_N &= r \log (N - r N_r) - \log N_r \\ &\quad - (r - 1) \log \{N - (r - q) N_r\} + \log \rho \end{aligned} \quad (15)$$

where  $N_1$  has been replaced by its value in terms of  $N_r$  and  $N$ . In considering a binary mixture in which the total number of sites is kept constant there is



effectively only one independent variable, and equation (15) can be integrated without difficulty; we get

$$\log g(N_1, N_r) = -\log N_r! - \log \frac{(N-rN_r)!}{N!} + \frac{r-1}{r-q} \log \frac{\{N-(r-q)N_r\}!}{N!} + N_r \log \rho \quad (16)$$

where the constant of integration has been determined by the fact that when  $N_r$  is zero,  $g(N_1, 0)$  is unity so that the right-hand member of this equation must vanish. Re-introducing the parameter  $N_1$  this gives

$$\log g(N_1, N_r) = -\log N_r! - \log N_1! + \frac{r-1}{r-q} \log (N_1+qN_r)! - \frac{q-1}{r-q} \log (N_1+rN_r)! + N_r \log \rho \quad (17.1)$$

or

$$g(N_1, N_r) = \frac{(N_1+rN_r)!}{N_1! N_r!} \left\{ \frac{(N_1+qN_r)!}{(N_1+rN_r)!} \right\}^{\frac{r-1}{r-q}} \rho^{N_r} \quad (17.2)$$

for the combinatory factor for a binary mixture of monomer and polymer molecules. This formula is valid whether the polymer molecules are simple chain, branched chain, or closed ring molecules.

For simple chain and branched chain molecules, these formulae take particularly simple forms by virtue of the value of  $q$  given by equation (1.1). Substituting this value in equations (17) we get

$$\log g(N_1, N_r) = -\log N_1! - \log N_r! + \frac{1}{2}z \log (N_1+qN_r)! - (\frac{1}{2}z-1) \log (N_1+rN_r)! + N_r \log \rho \quad (18.1)$$

and

$$g(N_1, N_r) = \frac{\{(N_1+qN_r)!\}^{\frac{1}{2}z}}{N_1! N_r! \{(N_1+rN_r)!\}^{\frac{1}{2}z-1}} \rho^{N_r} \quad (18.2)$$

This form, which is valid for a binary monomer-polymer mixture in which the polymer consists of simple chain or branched chain molecules, admits of an obvious generalization which will be deduced strictly below (equation (26)).

(c) Now consider a mixture of two types of polymer molecules,  $N_i$  of type  $i$  and  $N_j$  of type  $j$ , and suppose also that there are at least a small number,  $N_1$ , of monomer molecules. The monomer molecules are introduced as an intermediary in part of the mathematical deduction and later it is supposed that  $N_1$  vanishes;  $N_1$  need only be of the order of the greater of  $r_i$  or  $r_j$ .

Considering the replacement of a molecule  $i$  by  $r_i$  monomer molecules a partial differential equation analogous to equation (15) is obtained. Thus

$$\left[ \frac{\partial}{\partial N_i} \log g(N_1, N_i, N_j) \right]_{N, N_j} = r_i \log (N-r_iN_i-r_jN_j) - \log N_i - (r_i-1) \log \{N-(r_i-q_i)N_i-(r_j-q_j)N_j\} + \log \rho_i \quad (19.1)$$

where now the differentiation is carried out not only with the total number,  $N$ , of sites of the array constant, but also with  $N_j$  constant. Likewise, considering the replacement of a molecule  $j$  by  $r_j$  monomer molecules we get

$$\left[ \frac{\partial}{\partial N_j} \log g(N_1, N_i, N_j) \right]_{N, N_i} = r_j \log (N - r_i N_i - r_j N_j) - \log N_j \\ - (r_j - 1) \log \{N - (r_i - q_i)N_i - (r_j - q_j)N_j\} + \log \rho_j \\ \dots\dots\dots (19.2)$$

From these equations we obtain

$$\left[ \frac{\partial}{\partial N_i} \log g(N_1, N_i, N_j) \right]_{N, N_j} - \frac{r_i}{r_j} \left[ \frac{\partial}{\partial N_j} \log g(N_1, N_i, N_j) \right]_{N, N_i} \\ = -\log N_i + \frac{r_i}{r_j} \log N_j + \left(1 - \frac{r_i}{r_j}\right) \log \{N - (r_i - q_i)N_i - (r_j - q_j)N_j\} \\ + \log \rho_i - \frac{r_i}{r_j} \log \rho_j \dots\dots\dots (19.3)$$

Now, the right-hand member of equation (19.3) does not involve  $N_1$  explicitly; when  $N_1$  is zero the third term can be written more simply as  $\log (q_i N_i + q_j N_j)$ . Also, in the limit as  $N_1$  approaches zero, the left-hand member has the value

$$\left[ \frac{\partial}{\partial N_i} \log g(N_i, N_j) \right]_N$$

Thus, dispensing now with the monomer molecules, which were introduced merely to simplify the derivation of the required equation, we obtain, for a binary mixture of molecules  $i$  and  $j$ ,

$$\left[ \frac{\partial}{\partial N_i} \log g(N_i, N_j) \right]_N = -\log N_i + \frac{r_i}{r_j} \log \{(N - r_i N_i)/r_j\} \\ + \left(1 - \frac{r_i}{r_j}\right) \log \left\{ \frac{q_j}{r_j} N + \frac{q_j r_j - q_j r_i}{r_j} N_i \right\} \\ + \log \rho_i - \frac{r_i}{r_j} \log \rho_j \dots\dots\dots (20)$$

where now, since  $N_i$  is the only *independent* variable,  $N_j$  must be given its value as a function of  $N_i$  in all the terms of the right-hand member of equation (20). The integration of this equation is straightforward. We obtain

$$\log g(N_i, N_j) = -\log N_i! - \log \frac{\{(N - r_i N_i)/r_j\}!}{N!} \\ + \frac{r_j - r_i}{r_j q_i - r_i q_j} \log \frac{\{q_j N/r_j + (r_j q_i - r_i q_j)N_i/r_j\}!}{N!} \\ + N_i \log \rho_i + \left( \frac{N - r_i N_i}{r_j} \right) \log \rho_j \dots\dots\dots (21)$$

where the constant of integration has been determined by the fact that when  $r_i$  is unity, so that  $q_i$  and  $\rho_i$  are also unity, the right-hand member must reduce to

an equation identical in form with equation (16). Re-introducing the variable  $N_j$  we get

$$g(N_i, N_j) = \frac{(r_i N_i + r_j N_j)!}{N_i! N_j!} \left[ \frac{(q_i N_i + q_j N_j)!}{(r_i N_i + r_j N_j)!} \right]^{\frac{r_j - r_i}{r_j q_i - r_i q_j}} \rho_i^{N_i} \rho_j^{N_j} \dots \quad (22)$$

This formula is valid for a binary mixture of any kind of polymer molecules—simple chain, branched chain, or closed ring. If both components consist of either simple chain or branched chain molecules the power of the second factor on the right-hand side of this equation assumes a simpler form independent of  $N_i$ ,  $q_i$ ,  $r_j$ , and  $q_j$ , but not otherwise. When both components consist of simple chain or branched chain molecules, we get

$$g(N_i, N_j) = \frac{(r_i N_i + r_j N_j)!}{N_i! N_j!} \left[ \frac{(q_i N_i + q_j N_j)!}{(r_i N_i + r_j N_j)!} \right]^{\frac{1}{2}} \rho_i^{N_i} \rho_j^{N_j} \dots \quad (23)$$

(d) It requires no elaboration to see that the analysis by which equations (19) were obtained can be extended to cover the case of a mixture containing any number of components. As before, the total number of particles is kept fixed so that the number of independent variables is one less than the number of components. The number of a particular kind of monomer molecule, which if necessary can be introduced as an intermediary, to simplify the derivation of the equations, in the same way as in Section IV (c), is taken throughout as a function of the numbers of molecules of other kinds. Proceeding as in the previous section, one obtains a set of partial differential equations

$$g_k = \left[ \frac{\partial \log g(N_i)}{\partial N_k} \right]_N = -\log N_k - (r_k - 1) \log \{N - \sum_j (r_j - q_j) N_j\} \\ + r_k \log \{N - \sum_j r_j N_j\} + \log \rho_k, \quad (\text{all } k) \quad \dots \quad (24)$$

where the summations are carried out over the independent components. The conditions that this set of partial differential equations may be integrable is expressed by a set of equations

$$\frac{\partial g_k}{\partial N_i} = \frac{\partial g_i}{\partial N_k}, \quad (\text{all } i, k) \quad \dots \quad (25.1)$$

This requires that

$$r_k \frac{r_i}{N - \sum_j r_j N_j} - (r_k - 1) \frac{r_i - q_i}{N - \sum_j (r_j - q_j) N_j}$$

and the similar expression obtained by interchanging  $i$  and  $k$  shall be equal for all  $i$  and  $k$ . This condition reduces to

$$(r_k - 1)(r_i - q_i) = (r_i - 1)(r_k - q_k) \quad (\text{all } i, k) \quad \dots \quad (25.2)$$

which implies that  $(r_i - 1)/(r_i - q_i)$  shall be independent of  $i$  for all  $i$ . Reference to equations (1) shows that condition (23) is always satisfied provided only simple chain and branched chain molecules are present in the mixture. Integrating the set of partial differential equations in this case, and determining the constant

of integration by the fact that if all but any pair of components is absent, the combinatory factor must reduce to the forms already found for binary mixtures, equations (17) and (23) above, we get

$$\begin{aligned} \log g(N_i) = & -\sum_i \log N_i! + \frac{1}{2}z \log (\sum_i q_i N_i) \\ & -(\frac{1}{2}z-1) \log (\sum_i r_i N_i)! + \sum_i N_i \log \rho_i \dots\dots (26) \end{aligned}$$

This verifies the statement following equation (18).

(e) For closed ring molecules, on the other hand,

$$\frac{r_i-1}{r_i-q_i} = \frac{1}{2}z \frac{r_i-1}{r_i} \dots\dots\dots (27)$$

and the condition (25) for the integrability of the partial differential equations is not in general satisfied. It is only satisfied if there are only two components, or in case there are more than two components including a polymer consisting of closed ring molecules all the other components consist of monomer molecules. Thus, in the general case of a mixture of many components, containing more than one high polymer and one at least consisting of closed ring molecules, a value for  $g(N_i)$  cannot be obtained in a mathematically precise way.

It should be noted that the condition of integrability fails only for mixtures which contain *at least three components of which at least two are high polymers and at least one consists of closed ring molecules*. Thus the partial differential equations are integrable and mathematically precise formulae, for the combinatory factor can be obtained for (a) all binary mixtures whether the components consist of simple, simple chain, branched chain, or closed ring molecules; (b) mixtures of any number of components containing not more than one high polymer species, which can consist of closed ring molecules equally well as of simple chain or branched chain molecules (this covers the case of a polymer dissolved in a mixed solvent); and (c) mixtures of any number of components including more than one high polymer species provided these consist only of simple chain or branched chain molecules.

For practical purposes, however, a value for the combinatory factor can be obtained even in those cases in which a mathematically precise formula cannot be derived rigorously. Since the parameter  $r$  is very large (of the order of many thousands) then  $(r-1)/r$  differs insignificantly from unity. Thus with negligible error  $(r_i-1)/r_i$  in equation (27) could be replaced by unity, without introducing any error greater than that inherent in the model. To this approximation, equation (26) can also be applied to mixtures containing closed ring molecules.

(f) Since this analysis is for mixtures in which the energies of mixing of pairs of components are assumed zero the assumption of random mixing does not introduce any direct error. This is because for such a mixture the numbers of interactions of different kinds is irrelevant to the specification of the configuration of the assembly. However, it has been pointed out elsewhere\* that for mixtures in which the molecules of any component occupy more than one site on the

\* Ref. (9), closing paragraph of Ch. II, compare Section III (a) above.

quasi-crystalline array the number of possible configurations of molecules still depends on parameters which refer to pairs of sites, the  $x_{ii}$  of equation (2.2). These parameters define the group of sites which are occupied by a polymer molecule. This implies that as a result of the flexibility of the polymer molecules a segment of one molecule may be competing with a segment of another molecule, or with a distant segment of the same molecule, for a site which is a closest neighbour of a site occupied by another segment of the latter. This is not taken into account in the analysis so that there is a concealed error in the definition of complete randomness (cf. Guggenheim, 6, Section 9). Its effect, which is difficult to assess, is likely to be small, and in any case no greater than the degree of approximation inherent in the physical model itself.

## V. REFERENCES

- (1) FLORY, P. J.—*J. Chem. Phys.* **9** : 660 (1941); *Ibid.* **10** : 51 (1942).
- (2) HUGGINS, M. L.—*Ibid.* **9** : 440 (1941); *J. Phys. Chem.* **46** : 15 (1942); *Ann. N.Y. Acad. Sci.* **43** : 1 (1942).
- (3) MÜNSTER, A.—*Kolloidzshr.* **105** : 1 (1943); *Z. Naturforsch.* **1** : 311 (1946); *Ibid.* **2** : 272, 284 (1947).
- (4) FOWLER, R. H., and RUSHBROOKE, G. S.—*Trans. Faraday Soc.* **33** : 1272 (1937).
- (5) MILLER, A. R.—*Proc. Camb. Phil. Soc.* **39** : 54 (1943).
- (6) GUGGENHEIM, E. A.—*Proc. Roy. Soc. A* **183** : 203 (1944).
- (7) MILLER, A. R.—*Proc. Camb. Phil. Soc.* **38** : 109 (1942).
- (8) BETHE, H. A.—*Proc. Roy. Soc. A* **150** : 552 (1935).
- (9) MILLER, A. R.—“The Theory of Solutions of High Polymers.” (Clarendon Press : Oxford, 1948.)
- (10) GINGRICH, N. S.—*Rev. Mod. Phys.* **15** : 90 (1943).
- (11) FINBAK, C.—*Avh. Norske VidenskAkad.* No. 10 (1945).
- (12) MILLER, A. R.—*Proc. Camb. Phil. Soc.* **43** : 422 (1947).
- (13) FOWLER, R. H.—“Statistical Mechanics.” 2nd Ed. (Cambridge, 1936.)
- (14) MILLER, A. R.—*J. Chem. Phys.* **15** : 513 (1947).



## N-ACYL DERIVATIVES OF AROMATIC KETIMINES

By J. E. BANFIELD,\* G. M. BROWN,\* F. H. DAVEY,\* W. DAVIES,\*  
and T. H. RAMSAY\*

[Manuscript received June 1, 1948]

### Summary

It has been found that N-acyl aromatic ketimines in general have the property of combining with certain compounds with active hydrogen, such as alcohols, amines, and mercaptans. The reaction takes place at room temperatures and is most pronounced when the acyl group is aliphatic. This property of acyl ketimines can be weakened or prevented by steric hindrance.

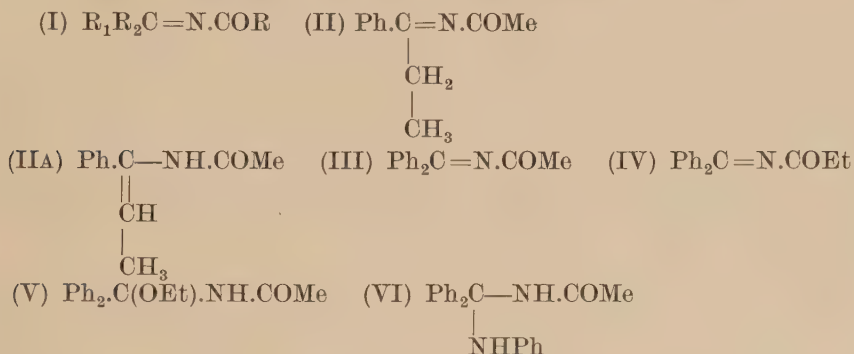
### I. INTRODUCTION†

There are several classes of compounds containing the azomethine group ( $C=N$ -) which can form addition compounds with alcohols and/or amines. Among the less obscure of these are the common reagents isocyanates ( $R.N=CO$ ), mustard oils ( $R.N=CS$ ), and the unstable acetylisocyanate  $CH_3.CO.N=CO(1)$ . In these compounds the  $C=N$  group is activated by a suitably disposed atom or group in the molecule. The structure of N-acyl ketimines (I) indicates that they likewise should possess abnormal additive powers, but the scientific literature of the few known members of this series does not record any additive properties at all. For example, N-benzoyldiphenylketimine,  $Ph_2C=NCOPh$ , was purified by crystallization from ethyl alcohol (Reddelien and Danilof,2). This apparent lack of additive power is also shown by the benzoyl and acetyl derivatives of aryl aliphatic ketimines, e.g. N-acetylphenylethylketimine (II), which for their purification require digestion with alcohol (Moureu and Mignonac,3). Nevertheless, the results of attempts by the present authors to prepare  $\alpha$ -chloracetyldiphenylketimine,  $Ph_2C=N.COCH_2Cl$ , in a manner similar to the preparation of (II), could only be explained by the addition of alcohol. Digestion with different alcohols gave different crystalline products the analysis of which showed that an alcohol had combined with  $\alpha$ -chloracetyldiphenylketimine. A thorough investigation of the problem has now shown that the addition of compounds with active hydrogen such as alcohols and amines is in fact a characteristic property of N-acyl aromatic ketimines of the general formula  $ArAr'C=N.COR$ . This property can in some cases be restricted by steric hindrance and other factors (p. 333).

\* Organic Chemistry Laboratory, University of Melbourne.

† Further information on the subject matter of this paper can be obtained from the theses for the degree of M.Sc. in the University of Melbourne, by F. H. Davey, 1945, G. M. Brown, 1946, and J. E. Banfield, 1948.

The best method of preparation of the N-acyl ketimines varies with the reactivity and ease of crystallization of the particular member of the series which is required. The relatively inert N-benzoyldiphenylketimine, contrary to the experience of Smith and Bergstrom(4), is made by the benzoylation of diphenylketimine followed by crystallization(2). N-acetylfluorenoneimine(5) is also prepared by direct acylation, which is the best way to make the easily crystallized *N-propionyldiphenylketimine* (IV) m.p. 79° C. *N-Acetyldiphenylketimine* (III), which has the low melting point of 45° C., can be made by direct acetylation of the ketimine followed by distillation under reduced pressure, and a good yield of pure material is also obtained by the pyrolysis of the ethyl alcohol derivative made from the crude acetylation product. In case of doubt or



difficulty the best way to make N-acetyl derivatives of ketimines is to pass ketene into a cold ethereal solution of the appropriate ketimine, and this method has been used to prepare (III), and also *N-acetylphenyl- $\alpha$ -naphthylketimine*, *N-acetylphenyl- $\beta$ -naphthylketimine*, and *N-acetyl- $\alpha\alpha$ -dinaphthylketimine*, and, in chloroform, N-acetylauramine.

The previously described N-benzoyldiphenylketimine slowly combines with alcohol at room temperatures, but the additive property of the diphenylketimine derivatives is much more pronounced when the acyl group is aliphatic. Thus N-acetyldiphenylketimine (III) readily combines with ethyl alcohol to form the *derivative* (V) and with aniline to form (VI). An approximate comparison of the rate of combination with ethyl alcohol in petroleum ether shows that (III) is more reactive than phenylisothiocyanate but rather less reactive than phenylisocyanate. Like these reagents it does not readily form derivatives with secondary alcohols, and not at all with tertiary alcohols or phenols. The method of combination usually adopted is to allow the two reactants in petroleum ether to stand at room temperature until the crystals cease to be deposited. In order to make fair comparisons, when alcohols are used, no basic catalyst has been added in reactions of which the time has been taken. N-Propionyldiphenylketimine (IV), owing to its easy preparation, is a more convenient reagent than (III), and an account of some of its derivatives is given in tabular form. The alcohol and amine derivatives of both (III) and (IV) all have melting points greater than 95° C.

Typical N-acyl aromatic ketimines such as (III) and (IV) react at ordinary temperatures with a variety of compounds containing active hydrogen. Ammonia, primary amines such as ethylamine and aniline, and some secondary amines such as piperidine readily form compounds which can usually be recrystallized from petroleum ether or benzene. Amino esters such as glycine and alanine ethyl esters also form *derivatives* of sharp melting point, though amino acids themselves give negative results. Compounds with the  $\text{-SH}$  group such as ethyl and benzyl mercaptans and thiophenol also readily yield pure crystalline *derivatives* involving one molecule of each reactant. Primary aliphatic and mixed aliphatic aromatic alcohols such as benzyl carbinol more slowly form crystalline derivatives, though here the melting points are not always sharp. In the purification of the alcohol derivatives prolonged boiling with a high boiling solvent is to be avoided otherwise partial fission may occur, and the attempted purification may lead to a lowering but not to a raising of the melting point. In some cases, e.g. with normal amyl alcohol, merely washing the precipitated derivative with petroleum ether is recommended. Wet alcohols give products of lower melting points than those obtained from anhydrous alcohols. Hydrolysis of (III) and (IV) by aqueous hydrochloric acid yields benzophenone, but dry hydrogen chloride converts them into diphenylketimine hydrochloride.

The addition compounds of alcohols and amines with acyldiphenylketimines slowly decompose at or about their melting points, and rapidly on pyrolysis, which is effected by attempted distillation of the dry compound. Data of the results, when derivatives of N-propionyldiphenylketimine (IV) are used, are given in Table 1. Pyrolysis of the alcohol derivatives usually regenerates the alcohol and the original acyl ketimine. The *benzyl mercaptan derivative* of (IV) behaves similarly, though the mercaptan is isolated as the disulphide, probably as a result of atmospheric oxidation.

The pyrolysis of amine derivatives of (IV) can proceed in two ways. *Reaction (A)*,  $\text{Ph}_2\text{C}(\text{NR}_1\text{R}_2).\text{NHCOEt} \rightarrow \text{Ph}_2\text{C}=\text{NCOEt} + \text{HNR}_1\text{R}_2$ , is similar to the one described above for the alcohol derivatives of (IV), and is the reversal of the process of formation. This has been realized with some derivatives of secondary amines, such as piperidine and ethylene imine. Incidentally, it is noteworthy that in this experiment no fission of the ethylene imine ring itself takes place, whereas it is known that the pyrolysis of benzoylethylene imine gives 2-phenyloxazoline(6).

The second pyrolysis of amine derivatives of (IV) is the conversion of the compound into propionamide and a substituted diphenylketimine, the "anil" of benzophenone.

*Reaction (B)*,  $\text{Ph}_2\text{C}(\text{NHR}).\text{NHCOEt} \rightarrow \text{NH}_2\text{COEt} + \text{Ph}_2\text{C}=\text{NR}$ , has been realized for primary amines with the addition products of (IV) and aniline, ethylamine, and benzylamine, and both reactions (A) and (B) can apparently occur with the ammonia derivative of (IV).

There is another type of *reaction (C)* which has not been closely studied. The pyrolysis of the *ethanol derivative* of N-phenylacetyldiphenylketimine gives



some phenylacetamide. Similarly when solutions of N-benzoyldiphenylketimine and (a) piperidine and (b) ethanol are refluxed for a long time, benzamide is formed in both cases. This reaction awaits further investigation.

The fact that the alcohol, mercaptan, and secondary amine derivatives of acyl ketimines are smoothly converted by heat into their original constituents, may have some importance in the production of polymeric materials. It recalls the same property common to phenolic and some other derivatives of isocyanates of having their additive power "masked" until the particular moment when it is required. Heat, in the presence of a relatively non-volatile dihydric alcohol or diamine, regenerates the original di-isocyanate which then forms a polymer with the added alcohol or amine(7).

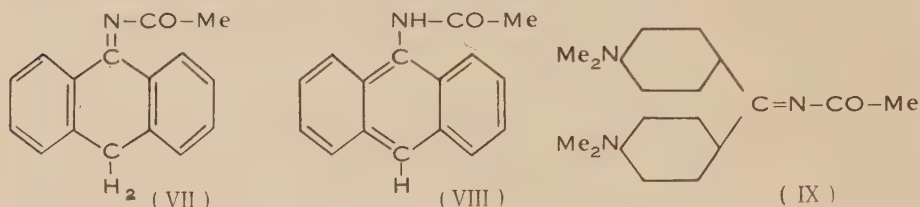
The N-acyl ketimines which add alcohols or amines or both are acetyl, chloroacetyl, propionyl, *n*-butyryl, and benzoyldiphenylketimines, also N-acetyl-fluorenoneimine and N-acetylphenyl- $\beta$ -naphthylketimine. In addition,  $\beta$ -chloropropionyl and phenylacetyldiphenylketimines, which have not been isolated in a pure condition, give pure *ethanol derivatives*.

The conventional explanation of the additive power of N-acyl ketimines is that the electrophilic acyl group imposes electrophilic character on the carbon atom doubly bound to the nitrogen atom. The  $\text{:C=N-}$  group then contains an incipiently polarized double bond comparable to that in benzylidene malonic ester, which also forms an additive compound with one molecule of aniline. The acyl group in N-acyl ketimines is clearly responsible for their unusual additive power, as can be seen by comparing them with the closely related Schiff's bases, which(8) will combine with strongly polarized molecules such as Grignard reagents and sodium bisulphite, but apparently not with alcohols or amines. Moreover, activation of the azomethine group in aromatic ketimines can be effected by other electrophilic groups than acyl; for example, it is now found that *N-carbethoxydiphenylketimine* and *N-carbethoxy-o-tolylphenylketimine* both combine with aniline, though the reaction is relatively slow with the *o*-tolylketimine derivative.

The following ketimine derivatives might therefore at first sight be expected to combine with amines and alcohols: *N-acetylphenyl- $\alpha$ -naphthylketimine*, *N-acetyl- $\alpha\alpha$ -dinaphthylketimine*, *NN'-di(diphenylmethylene)urea* (i.e.  $\text{Ph}_2\text{C=N-CO-N=CPh}_2$ ), and *N-acetylphenylethylketimine* (II). All these compounds, however, are inert. Steric hindrance is no doubt the reason for inactivity of the  $\alpha$ -naphthylketimines, as can be seen by comparing their full structural formulae with that of the sterically unhindered *N-acetylphenyl- $\beta$ -naphthylketimine*, which is active. In the *NN'-di(diphenylmethylene)urea* it is, on further reflection, not surprising that the effect of the one  $\text{-CO-}$  group is small when shared by two  $\text{-C=N-}$  groups.

*N-Acetylphenylethylketimine* (II), however, owes its inertia to its ability to exist in the tautomeric form, (IIA), which seems to express its properties better than does formula (II). For example, the compound reacts immediately with bromine and with aqueous permanganate, the latter reagent producing acetaldehyde. Methyl magnesium iodide at once causes the evolution of methane,

indicating the presence of an active hydrogen atom, which is contained in (IIA) but not in (II). No doubt the fact that acetylanthronimine (VII) does not react with alcohols and amines under the conditions successful with acetylfluorenoneimine is due to its existence in the tautomeric 9-acetaminoanthracene form (VIII), a change analogous to the formation of (IIA) from (II). The likelihood of such changes in anthronimine types is shown by Kaufler and Suchanek(9).



Some difficulty in interpretation of results is associated with N-acetylauramine (IX) which has been described (m.p.  $221^{\circ}\text{C.}$ ) by Semper(10), who prepared it from auramine base and acetic anhydride followed by treatment with alcohol. A repetition of his process gave a mixture m.p.  $140\text{--}174^{\circ}\text{C.}$ , whereas the use of acetic anhydride in a more direct manner, and also the action of ketene on the base, gave two identical products m.p.  $109\text{--}111^{\circ}\text{C.}$  This compound gives the striking colour reactions described by Semper for his product m.p.  $221^{\circ}\text{C.}$ , so apparently N-acetylauramine can exist in two forms. The one m.p.  $109\text{--}111^{\circ}\text{C.}$  shows a slight tendency to react with aniline but not with alcohol. The N-benzoyl derivative of auramine made after Finck and Schwimmer(11) has essentially the properties they ascribe to their compound, but does not readily combine with alcohols or amines. These and other doubtful cases are under investigation.

## II. EXPERIMENTAL

### (i) *Preparation of N-Acetyldiphenylketimine (III)*

When molar quantities of acetyl chloride and diphenylketimine are heated with excess of dry pyridine for four hours on a water-bath and the benzene extract distilled, a low yield of acetyldiphenylketimine is obtained. The following, however, are all preparative methods for this and other acyl ketimines.

(a) Ketene is passed into diphenylketimine (33 g.) in pure ether (50 ml.) at  $0^{\circ}\text{C.}$  until an excess of ketene is indicated by its pungent odour. The ether is removed, and 28 g. of the crude acetyldiphenylketimine, b.p.  $168\text{--}172^{\circ}\text{C./1 mm.}$ , are obtained; after redistillation, this crystallizes readily when seeded. The physical constants are R.I.  $n_{\text{D}}^{24} 1.6006$ , b.p.  $169\text{--}171^{\circ}\text{C./1 mm.}$ , and it crystallizes in characteristic rosettes from cold petroleum ether (b.p.  $40\text{--}50^{\circ}\text{C.}$ ), m.p.  $44\text{--}45^{\circ}\text{C.}$  (Found: N, 6.7%;  $\text{C}_{15}\text{H}_{13}\text{ON}$  requires N, 6.3%).



(b) Diphenylketimine (5 g.), acetic anhydride (3 ml.) and anhydrous pyridine (15 ml.) are refluxed for six hours, and the product when fractionated gives 3.8 g. of (III), b.p. 167-171° C./1 mm.

(c) The above quantities, with 15 ml. of benzene substituted for the pyridine, on refluxation for eight hours yield 4.0 g. of acetyldiphenylketimine, b.p. 168-171° C./1 mm., R.I.  $n_D^{24}$  1.6007.

(d) Benzonitrile (9.3 g.; 0.75 mole) is added to the Grignard reagent from magnesium (2.8 g.) and bromobenzene (20 g.), the total volume of ether being about 250 ml. After refluxation for half an hour, the crystalline product is washed with dry ether, then dispersed as much as possible in ether, and acetyl chloride (12 g.) in ether carefully added. When the violent reaction has subsided, the product is refluxed for five minutes, washed free from excess acetyl chloride with ether, and the residue refluxed for an hour with absolute alcohol (60 ml.), cooled and stirred with ice water for half an hour. The greenish solid is well drained, and this ethanol-N-acetyldiphenylketimine crystallizes from benzene and petroleum ether (needles, m.p. 166-168° C.), yield 3.5 g. It is pyrolysed and loses its alcohol at 210° C., and the residue on distillation (208-210° C./13 mm.) solidifies in the cold and crystallizes from 30-60° C. petroleum ether, m.p. 44-45° C. (Found: C, 80.55; H, 5.72%.  $C_{15}H_{13}ON$  requires C, 80.7; H, 5.83%.) The yield is over 80% in the actual pyrolysis of the alcohol compound, m.p. 166-168° C.

(e) The ether in the condensation of phenyl magnesium bromide from 3.6 g. magnesium and benzonitrile is replaced by benzene, and the whole refluxed with acetic anhydride (8 ml.) for quarter of an hour. The residue after cooling is well washed with benzene, and the combined benzene filtrates give a satisfactory yield of crystalline (III) after distillation.

Bromine in carbon tetrachloride is not decolorized by (III), which, however, in benzene solution reacts with hydrogen chloride to form a precipitate of diphenylketimine hydrochloride. (Found: C, 71.65; H, 5.7%.  $C_{13}H_{12}NCl$  requires C, 71.7; H, 5.5%.) Excess of acetyl chloride in benzene yields a colourless compound (m.p. 82-114° C.). Boiling dilute hydrochloric acid converts (III) into benzophenone.

#### (ii) *Preparation of Amine Derivatives of (III) and (IV)*

The acyl ketimine in 20 times its weight of petroleum ether is mixed with a slight excess over a molar quantity of the amine, and the solution allowed to remain at room temperature until precipitation appears complete. During crystallization of the product, prolonged boiling with solvents must be avoided, but rapid crystallization from petroleum ether (80-100° C.), in which it is sparingly soluble, or solution in warm benzene followed by precipitation with petroleum ether can be used. The following derivatives of (III) are obtained in over 80% yield by interaction of the reactants for several hours.

*Aniline-N-acetyldiphenylketimine*, m.p. 159-161° C. (Found: C, 79.9; H, 6.53%.  $C_{21}H_{20}ON_2$  requires C, 79.8; H, 6.33%) is converted into benzo-phenone anil, m.p. 113.5° C., when refluxed for 12 hours with petroleum ether (b.p. 120-135° C.).

*Ethylamine-N-acetyldiphenylketimine*, m.p. 117-119° C. (Found: C, 76.1; H, 7.6%.  $C_{17}H_{20}ON_2$  requires C, 76.1; H, 7.5%.)

*Phenylhydrazine-N-acetyldiphenylketimine* has m.p. 123-125° C. *Piperidine-N-acetyldiphenylketimine*, produced in a few seconds from the reactants, has m.p. 96-99° C. (Found: C, 78.2; H, 7.9%.  $C_{20}H_{24}ON_2$  requires C, 78.0; H, 7.8%.)

### *Ammonia Derivatives*

(a) Ammonia gas is passed into a dry benzene solution of acetyldiphenylketimine, and allowed to stand for six hours; and the white crystalline solid formed recrystallized from benzene, m.p. 117.5-118.5° C. (Found: C, 75.45; H, 6.9%.  $C_{15}H_{16}N_2O$  requires C, 75.05; H, 6.7%.)

(b) A second derivative of acetyldiphenylketimine with ammonia is obtained by adding a small amount of ammonium hydroxide (sp. gr. = 0.880) to its alcoholic solution. Recrystallized from benzene, it has m.p. 143-145° C. (Found: C, 75.1; H, 6.85%.  $C_{15}H_{16}N_2O$  requires C, 75.05; H, 6.68%.)

### (iii) *Alcohol Derivatives of N-Acetyldiphenylketimine (III)*

These are made by combining equimolecular quantities of the dry alcohol with (III) in petroleum ether at room temperatures; high temperatures have an adverse influence. The derivatives crystallize in characteristic matted colourless needles. In all cases the yield is over 70%, and the time for the reaction is indicated in parenthesis.

*Methanol-N-acetyldiphenylketimine* (2 days), m.p. 158° C. When slowly heated, begins to soften at 141° C. (Found: C, 75.3; H, 6.7%.  $C_{18}H_{17}O_2N$  requires C, 75.2; H, 6.4%.)

*Ethanol-N-acetyldiphenylketimine* (1 day; begins to separate in 5 hours). Also made by treating the reaction product of benzonitrile and phenyl magnesium bromide with acetyl chloride and then digesting the product with ethyl alcohol. (Found: C, 75.6; H, 6.8; N, 5.3%.  $C_{17}H_{19}O_2N$  requires C, 75.9; H, 7.06; N, 5.2%.) Both preparations have m.p. and mixed m.p. 166-168° C.

*n-Propanol-N-acetyldiphenylketimine* (2 days), m.p. 147-149° C.

*n-Butanol-N-acetyldiphenylketimine* (3 days), m.p. 128-129° C. (Found: C, 76.25; H, 7.7%.  $C_{19}H_{23}O_2N$  requires C, 76.6; H, 7.75%.)

*n-Pentanol-N-acetyldiphenylketimine* (5 days), m.p. 109-112° C. Recrystallization lowered the m.p., and analysis was carried out on a specimen purified merely by washing with petroleum ether. (Found: N, 4.75%.  $C_{20}H_{25}O_2N$  requires N, 4.5%.) The *benzyl carbinol* derivative of (III) (1 hour) has m.p. 163-166° C.

(iv) *Preparation of N-Chloracetyldiphenylketimine*

Chloracetyl chloride (3.1 g.; 1 mole) is slowly added to diphenylketimine (10 g.; 2 moles) in dry benzene, with continual stirring. Heat is evolved, and diphenylketimine hydrochloride is immediately precipitated. The mixture is allowed to stand for  $1\frac{1}{2}$  hours, when the hydrochloride is filtered off, and the solvent removed from the filtrate by evaporation under low pressure. The residue, which soon solidifies in rosettes, m.p. 61-63° C., is crystallized from petroleum ether (b.p. 40-60° C.), in which it is moderately soluble. (Found: N, 5.46%.  $C_{15}H_{12}ONCl$  requires N, 5.44%.)

The *ethanol derivative*, made in benzene and recrystallized from benzene-petroleum ether mixture, is identical (m.p. 170-171° C.) (mixed m.p.) with the compound obtained by using chloracetyl chloride in process (d) (page 335). (Found: N, 4.6%.  $C_{17}H_{18}O_2NCl$  requires N, 4.6%.)

The *methanol derivative* has m.p. 155-156° C. (Found: N, 4.86; Cl, 11.9%.  $C_{16}H_{16}O_2NCl$  requires N, 4.84; Cl, 12.3%.)

The *ethanol derivative* of  $Ph_2C=NCOCH_2Br$  (from bromacetyl bromide as in (d) (page 335), but not itself isolated) has m.p. 161-161.5° C. The *ethanol derivative*  $Ph_2C(OEt)NCOCH_2I$ , m.p. 166-166.5° C. (Found: I, 33.4%.  $C_{17}H_{18}O_2NI$  requires I, 32.9%) is made by the action of NaI in acetone on the ethanol derivative (m.p. 170-171° C.) of N-chloracetyldiphenylketimine above.

(v) *N-Propionyldiphenylketimine (IV) and Related Compounds*

Diphenylketimine (6 ml.), propionic anhydride (4 ml.) are heated to boiling for three minutes, and after cooling petroleum ether (5 ml.) added. The large crystals (5.8 g.; 71% yield) rapidly formed, separate from cyclohexane in prisms up to 1 cm. wide, m.p. 78-79.5° C. (Found: C, 80.45, 80.6; H, 6.17, 6.14%.  $C_{16}H_{15}ON$  requires C, 80.5; H, 6.33%.) It is also readily prepared by boiling diphenylketimine hydrochloride (0.5 g.), anhydrous sodium propionate (0.35 g.) and propionic anhydride (1 ml.) for about two minutes.

*Normal N-butyryldiphenylketimine* (m.p. 74-75° C. from petroleum ether) is made in 65% yield by both the above methods. (Found: C, 81.1; H, 6.55%.  $C_{17}H_{17}ON$  requires C, 81.2; H, 6.7%.) It forms a crystalline derivative with aniline in petroleum ether (4 hours).

*Ethanol-N-β-chloropropionyldiphenylketimine* (m.p. 157° C. from benzene) is made by using β-chloropropionyl chloride in (d) (page 335), followed by digestion with ethanol.

*N-Propionyl-o-tolylphenylketimine* (massive square crystals m.p. 46-47° C.), made by heating the imine with propionic anhydride for  $\frac{1}{2}$  hour at 160-170° C., does not form a derivative in one week with aniline, or piperidine, or ethyl alcohol. (Found: N, 5.80%.  $C_{17}H_{17}ON$  requires N, 5.57%.)

(vi) *Attempted Preparation of Phenylacetyldiphenylketimine*

Two moles of diphenylketimine are mixed with one of phenylacetyl chloride in petroleum ether, the diphenylketimine hydrochloride filtered off, ethanol added to the filtrate, and the immediate precipitate of white solid recrystallized



from ethanol, m.p. 181.5-182.5° C. (Found: N, 4.18%.  $C_{22}H_{23}O_2N$  requires N, 4.06%.) It is also formed by the addition of ethanol to the reaction product of phenylacetic anhydride and diphenylketimine. It yields phenylacetamide on pyrolysis, and all attempts to isolate the free acyl imine failed.

Table 1 depicts the derivatives of N-propionyldiphenylketimine (IV).

N-Benzoyldiphenylketimine, as prepared by Reddelien and Danilof(2) (who crystallized it from ethanol), had m.p. 118° C., but their analysis (C, 85.0; H, 4.57%) was not in good agreement for the theoretical figures ( $C_{20}H_{15}ON$  requires C, 84.2; H, 5.27%), which they calculated in error as C, 84.8; H, 4.62%. Accordingly, the compound has been prepared in several ways without the use of alcohol. The product obtained by heating molecular quantities of diphenylketimine and benzoic anhydride in benzene for half an hour is distilled; the fraction 173-180° C./10 mm. solidifies and crystallizes from petroleum ether (b.p. 80-100° C.), m.p. 117-118° C., in over 70% yield. (Found: C, 83.8, 83.65; H, 5.66, 5.61%. Calculated: C, 84.2; H, 5.27%.) It can also be made without distillation, by using benzoic anhydride or benzoyl chloride with pyridine, followed by rapid washing with water; also by heating diphenylketimine hydrochloride, benzoic anhydride, and dry sodium benzoate for six hours and crystallizing rapidly from alcohol.

*Ethanol-N-benzoyldiphenylketimine.*—Benzoyldiphenylketimine is dissolved in absolute ethanol, next morning unchanged benzoyldiphenylketimine, m.p. 118° C., removed, and in two weeks the mother liquor deposits large plates, m.p. 158-159° C. (Found: C, 79.95; H, 6.43%.  $C_{22}H_{21}O_2N$  requires C, 79.8; H, 6.35%.) Similarly, *n*-butanol-*N*-benzoyldiphenylketimine (needles, m.p. 151.5 to 152.5° C.) is formed in two weeks at ordinary temperatures. (Found: N, 4.08%.  $C_{23}H_{25}O_2N$  requires N, 3.90%.)

Refluxing benzoyldiphenylketimine with ethyl or *n*-butyl alcohol in petroleum ether (80-100° C.) for five days gives benzamide (m.p. 126° C.).

Aniline and piperidine *N*-benzoyldiphenylketimines are not isolated on admixture, but heating brings about the formation of benzamide, indicating that the reaction (B) or (C) (p. 332) has probably occurred.

*N*-Acetylfluorenoneimine.—The intermediate formation of the picrate(5) of the imine is avoided by shaking, for 5 hours, 5 g. of the crimson stannous chloride complex from fluorenone oxime with a mixture of ammonia (33 ml. of sp. gr. 0.880) and ether (40 ml.). The aqueous layer is extracted with more ether, and the combined ether extracts, dried over  $Na_2SO_4$ , yield the orange imine (2.7 g.) pure enough for acetylation with acetic anhydride in benzene after Kliegl(5).

*Ethanol-N-acetylfluorenoneimine* (1 week), m.p. 135-135.5° C. (Found: C, 75.9; H, 7.1%.  $C_{17}H_{17}O_2N$  requires C, 76.3; H, 6.4%.)

*Aniline-N-acetylfluorenoneimine* (1 hour) from benzene (yellow), m.p. 142-143° C. (Found: C, 80.8; H, 6.59%.  $C_{21}H_{18}ON_2$  requires C, 80.3; H, 5.7%.)

TABLE I  
ADDITION PRODUCTS WITH N-PROPYLDIPHENYLKETIMINE (IV)

Compound	Melting Point (°C.)	C		H		N		Time for First Crystals	Pyrolysis Gives
		Fd.	Th.	Fd.	Th.	Fd.	Th.		
Ethanol ..	174-176	76.8	76.3	7.54	7.4			1 day	Ethanol + (IV)
n-Butanol ..	159-161							1 day	
Ethyl mercaptan ..	138-139	72.7	72.3	7.21	7.03			30 minutes	
Benzyl mercaptan ..	140-141	76.55	76.5	6.90	6.42			$\frac{1}{2}$ day	Dibenzyl disulphide + (IV)
Thiophenol ..	127	76.5	76.2	6.95	6.1			$\frac{1}{2}$ hour	
Ammonia ..	109-111					11.17	11.02	$\frac{1}{2}$ hour after gas passed into hot solution	Ammonia, diphenylketimine + (IV)
Piperidine ..	105-106	77.7	78.3	7.81	8.07			2 minutes	Piperidine + (IV)
Benzylamine ..	140-141					8.20	8.14	Few seconds	Propionamide + N-benzyl diphenylketimine
Methyl anthranilate ..	146-153					7.04	7.22	1 hour	
$\beta$ -Phenylethylamine ..	112					7.9	7.82	1 minute	Viscous oil. Did not solidify
Ethylamine ..	100-101							Immediately	Propionamide + N-ethyl benzophenone imine
Aniline ..	137-138					8.65	8.48	5 minutes	Propionamide + benzophenone anil
Ethylene imine ..	125							<1 minute	Ethylene imine + (IV)
$\alpha$ -Phenylethylamine ..	82-83	76.95	77.1	7.87	7.15			<2 minutes	(IV) + $\alpha$ -phenylethylamine
Morpholine ..	150-151					8.68	8.65	4 minutes	(IV), morpholine
Alanine ethyl ester ..	126-127					7.97	7.83	Few seconds	

No catalyst was used in the reactions shown, all of which took place in petroleum ether (b.p. 40-60° C.) except in ethanol and butanol, when no solvent was employed. Alkali catalyses the addition of alcohols enormously, e.g. a little NaOEt with ethanol or solid KOH with methanol causes crystallization of the alcohol derivative in less than ten minutes.

Pyrolysis in general is effected by heating in a distillation flask at 200° C. for quarter of an hour, followed by cooling to about 150° C., and then distillation at about 20 mm.

The pyrolysis products were almost colourless except with the piperidine and morpholine derivatives, where the residues in the distilling flask were dark.

The reaction of glycine and alanine and (IV) gives negative results, and the addition of NaOEt to their alcoholic solution merely gives the ethanol derivative of (IV). The corresponding ethyl esters, however, rapidly combine with (IV) when melted together or in petroleum ether. The *glycine ethyl ester derivative*, needles, m.p. 133-134° C. (Found : OEt, 13.3%;  $C_{10}H_{17}O_3N_2$  requires OEt, 13.2%), which can be hydrolysed to benzophenone, when pyrolysed at 210° C./20 mm. yields propionamide and *ethyl diphenylketiminoacetate*,  $Ph_2C=N.CH_2.COOEt$ , needles m.p. 54-55° C., from petroleum ether. (Found : N, 5.6; OEt, 16.7%;  $C_{17}H_{17}O_3N$  requires N, 5.2; OEt, 16.9%.)



*Piperidine-N-acetylfluorenoneimine* (5 minutes in benzene and petroleum ether), m.p. 146-149° C. (Found: C, 79.2; H, 7.25%.  $C_{20}H_{22}ON_2$  requires C, 78.4; H, 7.2%.)

*N-Acetylphenyl- $\alpha$ -naphthylketimine*, colourless needles, m.p. 88-88.5° C. from petroleum ether (b.p. 60-80° C.) (found: C, 83.75; H, 5.67%.  $C_{19}H_{15}ON$  requires C, 83.5; H, 6.5%) and *N-acetyl- $\alpha$ -dinaphthylketimine* (straw-coloured needles, m.p. 169-170° C., from benzene and petroleum ether mixture) are made in 95% yield in ether at 0° C. by the ketene process. (Found: C, 84.7; H, 5.35%.  $C_{23}H_{17}ON$  requires C, 85.5; H, 5.26%.) Neither of these  $\alpha$ -naphthyl derivatives reacts with ethanol or aniline in petroleum ether.

*N-Acetylphenyl- $\beta$ -naphthylketimine* (large cream-coloured crystals, m.p. 64-67° C.) is also made by the ketene process. (Found: N, 5.31%.  $C_{19}H_{15}ON$  requires N, 5.13%.) The *aniline derivative*, m.p. 162-164° C., is purified by washing with petroleum ether. (Found: C, 81.3; H, 6.2%.  $C_{25}H_{22}ON_2$  requires C, 81.9; H, 6.0%.) The *piperidine derivative* has m.p. 119° C. The parent *phenyl- $\beta$ -naphthylketimine* made from phenyl magnesium bromide and  $\beta$ -naphthonitrile crystallizes from petroleum ether (b.p. 60-80° C.) in pale yellow crystals, m.p. 47-49° C., b.p. 234-236° C./19 mm. (Found: N, 5.7%.  $C_{17}H_{13}N$  requires N, 6.1%.)

The *phenyl- $\beta$ -naphthylketimine hydrochloride* and hydroxylamine hydrochloride in pyridine give in a few seconds the known lower melting oxime (m.p. 157-159° C.) of Poccianti(12), who records m.p. 157° C.

The properties of phenyl- $\alpha$ -naphthylketimine, m.p. 66-68° C., b.p. 230° C./23 mm., made from phenyl magnesium bromide and  $\alpha$ -naphthonitrile, agree with those of the product (m.p. 68-69° C., b.p. 181.5° C./4.5 mm.) made from benzonitrile and  $\alpha$ -naphthyl magnesium bromide by Moureu and Mignonac(13).

$\alpha$ -Dinaphthylketimine, m.p. 88° C., is made after Tschitschibabin and Korjagin(14).

Acetylphenylethylketimine, m.p. 125° C., made by the method of Moureu and Mignonac(13) does not combine with amines and alcohols. The presence of the ethylenic link is shown by instantaneous decolorization of bromine in carbon tetrachloride solution. Aqueous potassium permanganate yields acetaldehyde—isolated from the steam distillate of the oxidation product as the *p*-nitrophenylhydrazone, m.p. 124-126° C. (red), proved by the method of mixed melting points. The presence of active hydrogen (-NH-) is shown by the evolution of methane from MeMgI.

*N-Acetylanthroneimine* or 9-acetylaminoanthracene, m.p. 273-274° C. (yellow), made by the method of Goldman(15), is inert to aniline, piperidine, and ethanol under the conditions which are successful with *N*-acetylfluorenoneimine.

#### (vii) *Derivatives of Diarylketimine-N-carboxylic Acids*

*Diphenylketimine-N-formyl chloride*.—To a solution of phosgene in dry benzene (50 ml. of 20% weight by volume solution) diphenylketimine (15 g.) is gradually added. The immediately formed diphenylketimine hydrochloride

is separated, and distillation under reduced pressure yields 5 g. of a fraction 140-143° C./1.5 mm.,  $n_D^{14}$  1.6013. It is a very unstable liquid; a specimen redistilled and sealed in a glass tube for analysis becomes yellow before the analysis can be carried out. It is extremely reactive; several drops on a watch-glass exposed to air become completely solid in an hour through conversion to diphenylketimine hydrochloride. This also is the sole solid product of the action of anhydrous formic acid. Piperidine, aniline, and methyl aniline in benzene all react with diphenylketimine-N-formyl chloride to give amorphous solids containing halogen. Finally, a pure derivative is obtained by mixing equivalent amounts of the formyl chloride derivative and diphenylketimine in benzene; the diphenylketimine hydrochloride is separated, the benzene removed under reduced pressure, and the red residue deposits crystals overnight. These are colourless, m.p. 103-103.5° C., when crystallized from a mixture of benzene and petroleum ether. (Found: N, 7.33%.  $C_{27}H_{20}ON_2$  requires N, 7.22%.) This *NN'*-di(phenylmethylene) urea does not combine with amines or alcohols.

*N*-Carbethoxydiphenylketimine is made from diphenylketimine (2 moles) and ethyl chloroformate (1 mole) in benzene. After filtration and removal of the benzene, only one large fraction, 130-160° C./1 mm., is obtained. This gradually forms large clusters of crystals, m.p. 55.5-56.5° C., when recrystallized from petroleum ether (b.p. 40-60° C.). (Found: N, 5.53%.  $C_{16}H_{15}O_2N$  requires N, 5.53%.)

*Aniline-N*-Carbethoxydiphenylketimine (2 days in petroleum ether required) has m.p. 129-132° C.

*N*-Carbethoxyphenyl-*o*-tolylketimine.—To the Grignard compound from magnesium (2.8 g.) and *o*-bromotoluene (22.2 g.) in ether (150 ml.), benzonitrile (9.3 g.) is added, and refluxed for one hour. Ethyl chloroformate (16.8 g.) in its own volume of ether is added, and after refluxing for six hours the product is poured into ice water, the ether layer extracted with sodium bicarbonate solution, dried and distilled. 16 g. (66% yield) of *N*-carbethoxyphenyl-*o*-tolylketimine, b.p. 235-240° C./40 mm., is obtained. (Found: C, 76.8; H, 6.0; N, 5.0%.  $C_{17}H_{17}NO_2$  requires C, 76.4; H, 6.4; N, 5.2%.) The *aniline derivative*, m.p. 113-114° C., required six days for its formation at ordinary temperature in petroleum ether. (Found: N, 7.8%.  $C_{23}H_{24}O_2N_2$  requires N, 7.8%.)

#### (viii) Acyl Derivatives of Auramine

Commercial auramine, which contains much Michler's ketone, is purified as follows: Auramine hydrochloride (150 g.) is suspended in dry chloroform (400 ml.) and saturated with dry ammonia. The ammonium chloride is filtered off, washed with hot chloroform, and the filtrate and washings evaporated under reduced pressure. On cooling, auramine crystallizes and is recrystallized from absolute ethanol, m.p. 134-136° C. (about 40 g.).

*N*-Benzoylauramine.—Equal amounts of auramine and benzoic anhydride in benzene are boiled for a few minutes, allowed to remain overnight, the benzene is distilled off, and the residual dark solid rubbed with cold aqueous sodium

hydroxide until it is yellow. Crystallization from acetone, methyl or ethyl alcohol gives the benzoyl derivative, m.p. 179° C., identical with that obtained by Finck and Schwimmer(11). It gave no methane with MeMgI, showing the solvents had not combined. The addition of ethyl alcohol, aniline, or piperidine does not take place after standing for months.

*N-Acetylauramine*.—A solution of 4 g. of auramine base in the minimum amount of chloroform is saturated with ketene. Excess chloroform is removed under reduced pressure and the deep red residue extracted with the minimum amount of petroleum ether (60-80° C.), in which it is not very soluble. Fine yellow crystals, which darken on exposure to light, are obtained, m.p. 109-111° C. (Found: C, 73.65; H, 7.75%.  $C_{19}H_{23}ON_3$  requires C, 73.79; H, 7.44%.) It is also prepared when one drop of concentrated sulphuric acid is added to a cold solution of auramine base in excess of acetic anhydride. After remaining at room temperature for 12 hours, the solution is diluted with water and made slightly ammoniacal with ammonia, and the yellow precipitate recrystallized from petroleum ether, m.p. 109-111° C. Both the above methods give good yields.

A repetition of the process of Semper(10) gives products with m.p. from 140° to 174° C. It is noteworthy that Grandmougin and Favre-Ambrumyan(16), who use acetylauramine and refer to Semper's paper, give no details of the compound itself. As the present compound, m.p. 109-111° C., gives the striking colour reactions of Semper's compound, m.p. 221° C., *N*-acetylauramine evidently exists in two different crystalline forms.

*N*-acetylauramine, m.p. 109-111° C., which does not combine with ethyl alcohol, shows a slight tendency to react with aniline. After three weeks in petroleum ether at room temperatures, this gave a small quantity of yellow needles, m.p. 203-205° C., but a repetition on a larger scale produced after three months "auramine anil", m.p. 170° C., identified by mixed melting point.

### III. ACKNOWLEDGMENT

Microanalyses in connection with this work were carried out by N. L. Lottkowitz.

### IV. REFERENCES

- (1) SCHOLL, R.—*Ber. dtsh. chem. Ges.* **23**: 3510 (1890).
- (2) REDDELIEN, G., and DANILOF, H.—*Ibid.* **54**: 3139 (1921).
- (3) MOUREU, C., and MIGNONAC, G.—*Ann. Chim.* (9) **14**: 322 (1920).
- (4) SMITH, G. E. P., and BERGSTROM, F. W.—*J. Amer. Chem. Soc.* **56**: 2096 (1934).
- (5) KIEGL, A.—*Ber. dtsh. chem. Ges.* **43**: 2496 (1910).
- (6) GABRIEL, S., and STELZNER, R.—*Ibid.* **28**: 2933 (1895).
- (7) BAYER, O.—Polyurethanes. Combined Intelligence Objectives Sub-committee, C.I.O.S. Rep. No. XXIX, pp. 10-15 (Feb. 12, 1946).
- (8) GILMAN, H.—"Organic Chemistry", 2nd Ed., p. 659. (Wiley and Sons, 1943.)
- (9) KAUFLEDER, F., and SUCHANNEK, W.—*Ber. dtsh. chem. Ges.* **40**: 518 (1907).
- (10) SEMPER, L.—*Liebigs Ann.* **381**: 253 (1911).
- (11) FINCK and SCHWIMMER, O.—*J. prakt. Chem.* (2) **50**: 432 (1894).
- (12) POCCIANI, P.—*Brit. Abstr.* **108** (I): 822 (1915).
- (13) MOUREU, C., and MIGNONAC, G.—*C.R. Acad. Sci. Paris* **156**: 1806 (1913).
- (14) TSCHITSCHIBABIN, A. E., and KORJAGIN, S. I.—*Chem. Zbl.* **1914** (I): 1658 (1914).
- (15) GOLDMAN, F.—*Ber. dtsh. chem. Ges.* **23**: 2524 (1890).
- (16) GRANDMOUGIN, E., and FAVRE-AMBRUMYAN, S.—*Ibid.* **47**: 2129 (1914).



# THE THERMAL DECOMPOSITION OF SYNTHETIC AND NATURAL ALUNITE : AN INVESTIGATION BY X-RAY DIFFRACTION, ELECTRON DIFFRACTION, AND ELECTRON MICROSCOPE METHODS

By N. S. BAYLISS,\* J. M. COWLEY,† J. L. FARRANT,†  
and G. L. MILES‡

(Plates 1-4)

[*Manuscript received May 18, 1948*]

## Summary

In the thermal decomposition of synthetic alunite, as studied by X-ray diffraction, dehydration at about 500° C. is accompanied by the formation of dehydrated alum and  $\gamma$ -alumina. At 700–750° C.,  $\alpha$ -alumina predominates together with hexagonal potassium sulphate. At 850° C. there is some evidence of the formation of “potassium aluminate”. The calcination at 850° C. of a synthetic mixture of alunite and natro-alunite in the ratio 3 : 1 gives glaserite as the water-soluble phase.

The course of the thermal decomposition of natural alunite from Chandler, Western Australia, is similar except for important side reactions resulting from the presence of silica. The formation of an aluminium silicate (sillimanite or mullite), and of orthoclase or microcline was detected at temperatures as low as 750° C.

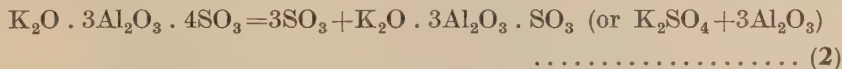
Electron micrographs are shown of natural alunite at various stages of decomposition.

## I. INTRODUCTION

The thermal decomposition of alunite ( $\text{K}_2\text{O} \cdot 3\text{Al}_2\text{O}_3 \cdot 4\text{SO}_3 \cdot 6\text{H}_2\text{O}$  or  $\text{KAl}_3(\text{SO}_4)_2(\text{OH})_6$ ) occurs in two main stages. The first is the loss of water vapour, the decomposition pressure reaching one atmosphere at 511° C.(1), and is represented by



The second stage, which begins at about 700° C., is the loss of sulphur trioxide (or its decomposition products sulphur dioxide and oxygen)



Potassium sulphate may be extracted by water leaching from the residue after the evolution of sulphur trioxide. In practice, even with pure alunite, the theoretical yield of potassium sulphate according to equation (2) is never obtained, and the yield decreases the higher the calcination temperature above about

\* Department of Chemistry, University of Western Australia.

† Division of Industrial Chemistry, C.S.I.R., Melbourne.

‡ State Alunite Industry (W.A.). Present address : Clare College, Cambridge, England.

800° C. The reason for this behaviour is doubtless the occurrence of side reactions in the solid, and the only previously reported attempt to use X-ray methods to elucidate the thermal decomposition and to identify the resulting solid phases is that of Fink, van Horn, and Pazour(2). Working with a pure natural alunite they identified dehydrated potash alum in the product formed by loss of steam at 500-700° C., and potassium sulphate and  $\alpha$ -alumina in the residue after loss of sulphur trioxide at 800° C. By heating to a considerably higher temperature (1400° C.) they were able to establish the occurrence of a reaction between potassium sulphate and alumina to give a new solid phase which they identified as a potassium aluminate of formula  $K_2O \cdot 10Al_2O_3$ .

In less pure forms of natural alunite, there is the added complication resulting from the possibility of gangue materials taking part in the high temperature reaction. The calcination and solubility study by Bayliss, Payne, and Pickering(3) on alunite from Chandler, Western Australia, showed that the optimum extraction of 86 per cent. of the theoretically available potassium sulphate was obtained after calcining at 700° C. Over-calcination of this material is particularly serious, since they found the yield of potassium sulphate to be reduced to about 40 per cent. after even a short calcination at 900° C. Since the aluminate formation reported by Fink, van Horn, and Pazour(2) did not set in until temperatures considerably above 1200° C., the present investigation was undertaken in an attempt to identify the reaction that was occurring and to determine the part played by the gangue material during the calcination of the Western Australian ore. Another point of interest was the form in which alumina occurred in the calcine, since Bayliss, Payne, and Pickering(3) found that its solubility in acids and in alkalis depended markedly on the calcination temperature.

The alunite deposit at Chandler occurs in the bed of Lake Campion, one of a chain of dry lakes belonging to the Lake Brown system(4). It is in the form of a plastic clayey material, and chemical analysis has shown that in addition to alunite as the major constituent, the material contains about 20 per cent. of silica as well as smaller amounts of natro-alunite, jarosite, and silicate minerals such as kaolinite. The deposit is impregnated with water containing dissolved sodium chloride together with some magnesium chloride. These chlorides are converted to sulphates during calcination. The resulting sodium sulphate together with the natro-alunite in the ore gives a calcine from which leaching yields potassium and sodium sulphates in a ratio that corresponds approximately to the composition of glaserite ( $K_3Na(SO_4)_2$ ). The principal crystal species obtained by leaching and crystallization is actually glaserite.

The investigation reported in this paper includes an X-ray study of the calcination of both pure synthetic alunite and Chandler alunite, together with an electron diffraction and electron microscope study of Chandler alunite. It was hoped to contribute to the solution of the following problems which have both theoretical and practical interest:

(a) Do potassium and sodium sulphates exist as separate crystal species in calcined Chandler alunite, or as glaserite?



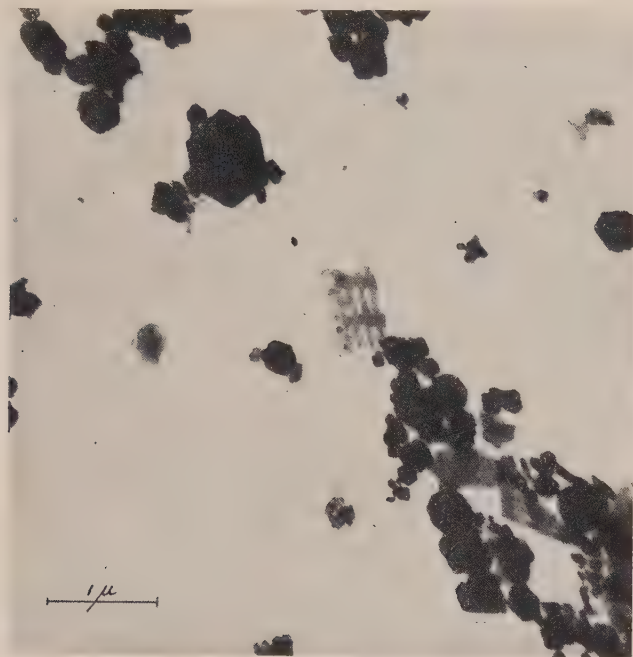


Fig. 1.

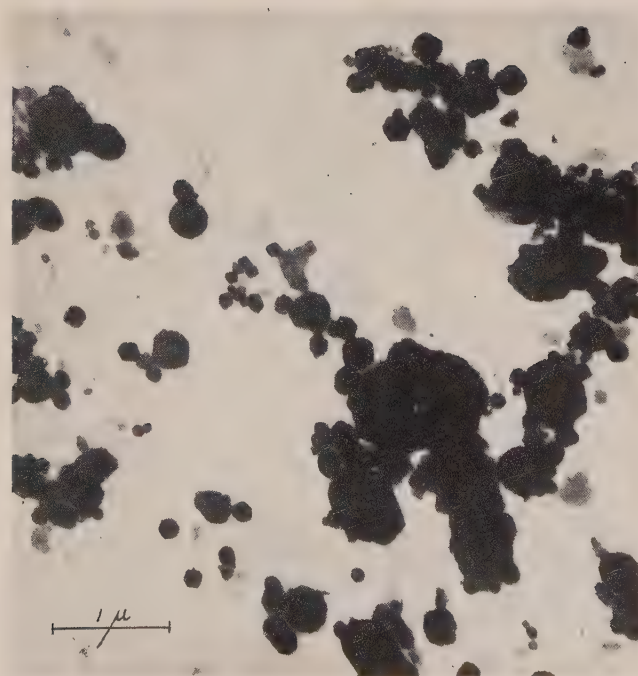


Fig. 2.



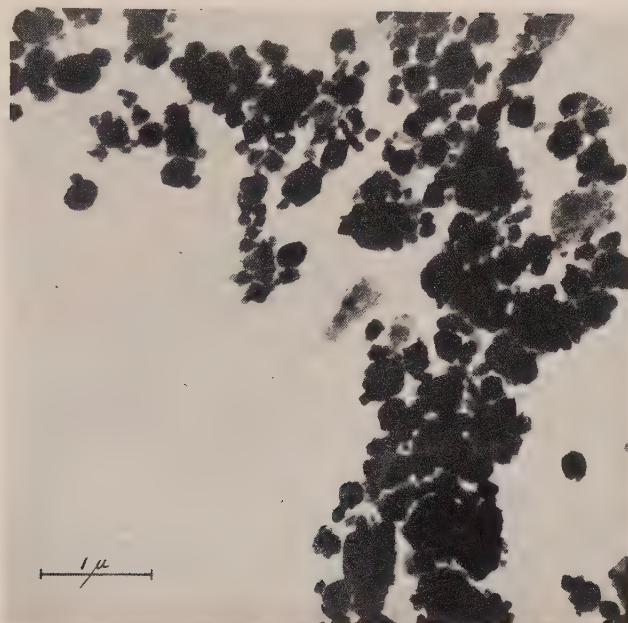


Fig. 1.

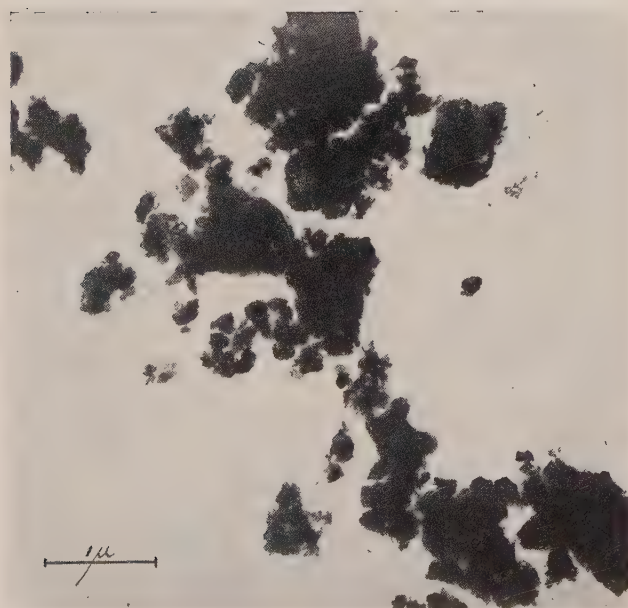


Fig. 2.



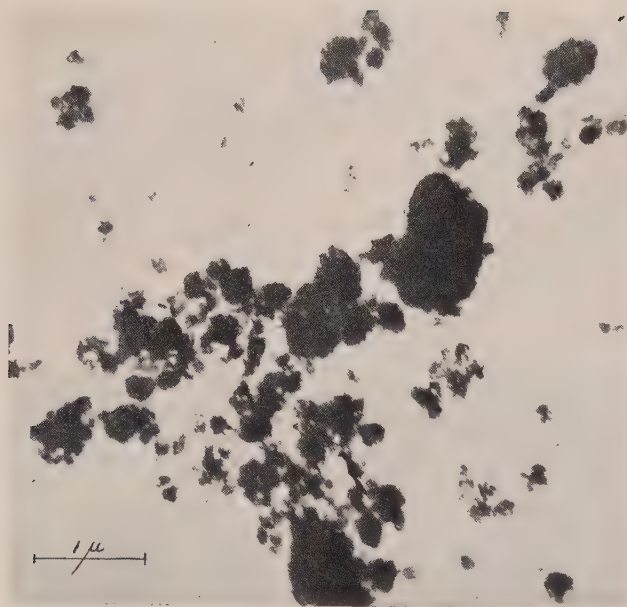


Fig. 1.

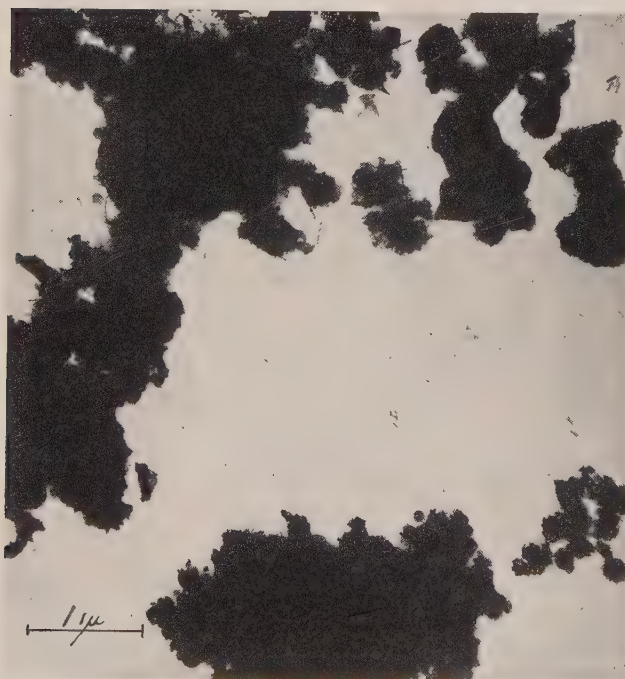
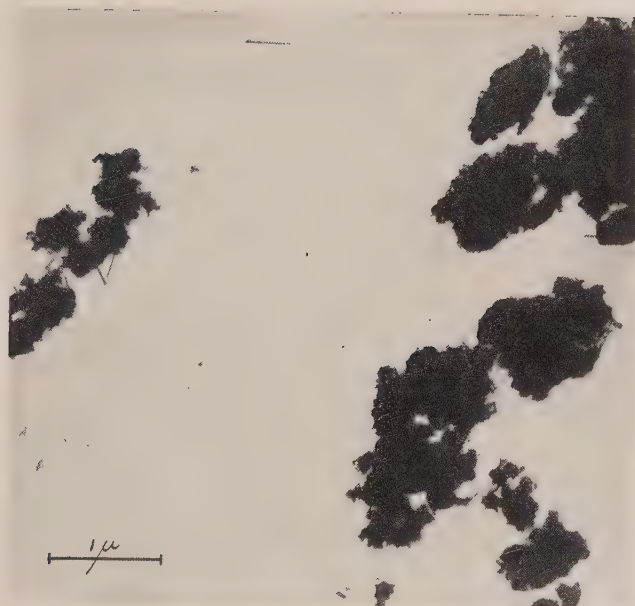


Fig. 2.







BAYLISS *et al.*.—THE THERMAL DECOMPOSITION OF SYNTHETIC AND NATURAL ALUNITE



(b) The crystalline form of the alumina produced at different calcination temperatures.

(c) The nature of the high temperature reaction that causes loss of water-soluble potash.

(d) The form of the silica in the Chandler material.

Ordinary microscopic examination(5) has shown that the ore consists almost entirely of particles of colloidal dimensions with only occasional larger crystals of quartz in an amount insufficient to account for the 20 per cent. given by chemical analysis.

## II. THE X-RAY STUDY OF SYNTHETIC ALUNITES (BY G.L.M.)

Alunite and natro-alunite ( $\text{NaAl}_3(\text{SO}_4)_2(\text{OH})_6$ ) were prepared in good yield by mixing portions of the corresponding alums with about twice their weight of water and heating at  $200^\circ\text{C}$ . in a small autoclave for 1-2 hours(6). The solid products were filtered and dried, and their identities were checked by their X-ray powder diffraction patterns. Samples of the pure (potassium) alunite and of a mixture of alunite and natro-alunite in the ratio 3 : 1 were calcined in a muffle furnace at temperatures between  $600^\circ$  and  $850^\circ\text{C}$ . The X-ray powder diffraction patterns were taken with copper radiation from a cold cathode X-ray tube and a camera of 5.7 cm. diameter.\* The patterns showed that potassium alunite calcined at  $600^\circ\text{C}$ . gave a solid residue containing  $\gamma$ -alumina and dehydrated alum. After calcining at  $750^\circ\text{C}$ ., the solid material gave the patterns of  $\alpha$ -alumina,  $\gamma$ -alumina, and hexagonal potassium sulphate. After calcination at  $850^\circ\text{C}$ ., four new lines appeared in the powder diffraction patterns that could not be ascribed to any of the solid phases mentioned above, but which correspond within our limits of error to four of the lines found by Fink, van Horn, and Pazour(2) and attributed by them to a phase of composition  $\text{K}_2\text{O} \cdot 10\text{Al}_2\text{O}_3$ . It is thus probable that some aluminate formation has occurred.

The mixture of synthetic alunite and natro-alunite was studied because of its bearing on the behaviour of the natural mineral. At  $600^\circ\text{C}$ . the patterns showed the presence of a mixture of dehydrated potassium and sodium alums together with  $\gamma$ -alumina and possibly a trace of hexagonal potassium sulphate. After calcination at  $850^\circ\text{C}$ . the glaserite pattern was found together with those of  $\alpha$ - and  $\gamma$ -alumina and some "aluminate". The patterns of the individual sodium and potassium sulphates were not present.

In view of the occurrence of reactions involving silica in natural alunite (Sections III, IV, and V), the following experiments were also performed :

(a) A mixture of finely crushed silica gel and  $\gamma$ -alumina was heated for 2 hours at  $875^\circ\text{C}$ . No new X-ray diffraction pattern was observed, indicating absence of any observable reaction under these conditions.

(b) A finely ground mixture of potassium sulphate and silica gel, containing a little iron oxide for its catalytic effect (see Section V), was heated for 2 hours at  $875^\circ\text{C}$ . The X-ray diffraction pattern provided no evidence of any silicate formation.

\* By courtesy of the Physics Department, University of Western Australia.

### III. THE X-RAY AND ELECTRON DIFFRACTION STUDIES OF CHANDLER ALUNITE (BY J.M.C. AND G.L.M.)

The work was carried out initially with the X-ray equipment mentioned in the previous section, and subsequently with the 14.3 cm. powder camera of the General Electric X.R.D. diffraction equipment and filtered cobalt radiation.\* The electron diffraction examination of some of the samples was made with the electron diffraction adapter of the RCA type EMU electron microscope.\*

The raw material from Chandler gave an X-ray diffraction pattern in which the alunite lines predominated. Other weak lines indicated the presence of  $\alpha$ -quartz, kaolin, sodium chloride, and the aluminosilicates orthoclase and microcline. The lines of  $\alpha$ -quartz were too weak to account for more than a small fraction of the silica indicated by chemical analysis. There was no sign of the lines of either natro-alunite or jarosite ( $\text{KFe}_3(\text{SO}_4)_2(\text{OH})_6$ ) showing that the sodium and the iron of the natural material must be incorporated by isomorphous replacement in the single alunite phase. Electron diffraction patterns showed the rings of alunite and kaolin. In an attempt to determine the form of the silica, the diffuse haloes of amorphous silica were sought but not found. Reflections from  $\alpha$ -quartz and the aluminosilicates were also present mainly in the form of a few sharp spots showing that they were present as comparatively large crystals.

After dehydration for 1 hour at 475° C., the X-ray pattern of alunite was weaker, and weak lines of  $\gamma$ -alumina appeared together with other weak lines probably resulting from partly broken-down alunite structures. After further dehydration at 600° C. only traces of the alunite pattern remained, and the principal X-ray diffraction lines were those of dehydrated alum ( $\text{KAl}(\text{SO}_4)_2$ ). Other weak lines were those of  $\alpha$ -quartz,  $\alpha$ -alumina, and one of the aluminium silicates, either mullite or sillimanite.

Natural alunite that had been calcined at 750° C. was examined before and after leaching with water so that the soluble and insoluble phases could be distinguished. The soluble material was mainly glaserite, with some hexagonal potassium sulphate. The principal constituent of the insoluble residue was an aluminium silicate, either sillimanite ( $\text{Al}_2\text{O}_3 \cdot \text{SiO}_2$ ) or mullite ( $3\text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2$ ). The powder patterns of these two materials are very similar, and the differences are not sufficient to permit a definite identification of the material in the calcine, although the observed spacings were in rather better agreement with the published data on sillimanite. Insoluble constituents, that were present in smaller amounts, were  $\alpha$ -quartz,  $\alpha$ -alumina,  $\gamma$ -alumina, and either microcline or orthoclase or both.

From the potash works at Chandler were received samples of calcined alunite that had been "over-calcined" at a temperature high enough to decrease the yield of soluble potash below its optimum value. The pattern from material that had been "slightly over-calcined" differed little from the results described in the previous paragraph, which correspond to optimum conditions. The only

\* By courtesy of the Division of Industrial Chemistry, C.S.I.R.



detectable difference was an increase in the intensity of the sillimanite lines. In a sample that had been "much over-calcined" the lines of both glaserite and sillimanite were weaker. New strong lines had appeared, some of which were due to  $\alpha$ -alumina. The others could not be identified unambiguously but almost certainly were given by aluminosilicates. Some agreement was found with the pattern of one type of orthoclase and with that of the sodium aluminium silicate  $\text{Na}_2\text{O} \cdot \text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2$ . The "aluminate" lines found in the pattern of calcined synthetic alunite could not be detected, although if they were present at low intensity they would have been masked by their proximity to the strong lines of the sillimanite pattern.

#### IV. ELECTRON MICROSCOPE EXAMINATION OF ALUNITE AND ITS DECOMPOSITION PRODUCTS (BY J.L.F.)

Owing to the presence of water-soluble constituents in some of the samples that were examined, it was necessary to use an organic liquid as a dispersing medium. Amyl alcohol was found to be satisfactory in all cases, and its use in turn demanded supporting films of polyvinyl formal which were prone to show long cracks and circular or oval holes, which, however, were readily recognized. Each sample was reduced to a fine state of subdivision in a steel mortar and the resulting powder was suspended in amyl alcohol. The suspension was shaken for five minutes, steel fragments were removed magnetically, and after then allowing the suspension to stand for three minutes to allow aggregates to settle, a sample was withdrawn from a point 0.5 cm. below the surface and was allowed to evaporate on the supporting film.

Micrographs were taken with the RCA type EMU electron microscope at a magnification of 7,500 diameters, and were enlarged optically to a magnification of 25,000 diameters.

Raw alunite from Chandler is shown in Plate 1, Figures 1 and 2. It is seen to consist of isometric crystals showing hexagonal, octagonal, and dodecagonal profiles ranging from 0.04 to 0.7 microns in diameter. These crystals are identified as alunite, which occurs as rhombohedra resembling cubes since the angle between the (100) and (010) faces is  $90^\circ 50'$ . In spite of the fact that Chandler alunite is known to contain silica, the only form of this substance that can be recognized in the micrographs consists of a few fragments of diatom skeletons (Plate 1, Fig. 1). The thin plates in Plate 1, Figures 1 and 2, are probably kaolinite or mica.

Plate 2, Figure 1, shows Chandler alunite that has been dehydrated for one hour at  $475^\circ\text{C}$ . The characteristic crystal habit of alunite persists, but is beginning to disintegrate. In Plate 2, Figure 2 (dehydration for 45 minutes at  $600^\circ\text{C}$ .) there are only the remaining vestiges of the alunite habit, and the sample consists of irregular aggregates with finely divided material adhering to their surfaces.

Chandler alunite that has been calcined under conditions to give the optimum yield of water-soluble potash ( $700$ - $750^\circ\text{C}$ .), and then leached, is shown in Plate 3, Figure 1. The micrograph shows no identifiable regular crystals; however,

the X-ray diffraction work on this sample shows that it contains a considerable amount of either sillimanite or mullite together with  $\alpha$ -alumina and other constituents mentioned previously.

Alunite, that has been "slightly over-calcined" with consequent loss of water-soluble potash, is shown before leaching in Plate 3, Figure 2, and after leaching in Plate 4. In addition to the irregular aggregates that appear to be characteristic of decomposed alunite, both samples show fine, rod-like prisms. Since these do not appear in Plate 3, Figure 1, it is natural to associate them with the reaction that causes the yield of potassium sulphate to decrease. The X-ray examination of this sample showed that the main difference from the sample shown in Plate 3, Figure 1, was an increase in the intensity of the sillimanite lines. However, the formation of more sillimanite does not of itself account for loss of potassium sulphate. The angles at the ends of the prisms in Plate 3, Figure 2, and Plate 4 were measured to be  $62 \pm 3^\circ$ , a value which is difficult to reconcile with the orthorhombic habit of sillimanite or mullite, although the (112) plane of sillimanite makes an angle of  $62^\circ$  with the long axis of its prisms. Prisms with angles of about  $62^\circ$  are also compatible with orthoclase, whose formation would lead to loss of water-soluble potash. Orthoclase was not recognized in the X-ray examination of this alunite sample, although it (or a similar aluminosilicate) was found in both the sample calcined under optimum conditions and in the sample described as "much over-calcined". The electron micrograph in Plate 3, Figure 1, of the sample calcined under optimum conditions shows no crystals that can be identified as orthoclase. These inconsistencies are probably related to crystal sizes and to sampling errors in view of the small samples required for both X-ray diffraction and the electron microscope.

## V. DISCUSSION

(a) *The calcination of pure synthetic alunite.*—The results reported in this paper are in general agreement with those of Fink, van Horn, and Pazour(2) with the exception that whereas the latter authors found alumina to be produced by calcination in its  $\alpha$ -form, we have been able to establish that the earliest form to appear at lower temperatures is  $\gamma$ -alumina, which is transformed to the  $\alpha$ -form as the calcination temperature is raised. The formation of  $\gamma$ -alumina is compatible with the ready solubility in acids of the alumina in alunite that has been calcined at moderate temperatures(3). The difference can be attributed to the long calcination times (13 to 48 hours) used by Fink, van Horn, and Pazour. Another difference in point of detail is the comparatively low temperature ( $850^\circ\text{C.}$ ) at which we found evidence for the formation of "potassium aluminate". From the viewpoint of the present problem, it is unfortunate that the characterization of compounds that can be termed potassium aluminate, and of their X-ray diffraction patterns, is so indefinite.

(b) *The calcination of natural alunite and the influence of silica.*—The calcination of natural alunite from Chandler (Sections III and IV) follows the same course as that of synthetic alunite (Section II), with the exception of the additional reactions that must be attributed to the silica in the natural material. In

common with the synthetic material, dehydrated alum,  $\gamma$ -alumina, and  $\alpha$ -alumina occur at successive stages in the decomposition. The water-soluble material has been identified as glaserite.

In addition to the above products, the appearance of sillimanite (or mullite), orthoclase, and possibly other aluminosilicates is obviously to be attributed to the silica in the natural ore. The remarkable feature is the low temperature at which they are formed. A survey of the literature available to us revealed no record of any reaction taking place between alumina and silica below  $1000^{\circ}\text{C}.$ , and the negative result of the experiment with a mixture of silica gel and  $\gamma$ -alumina (Section II) is confirmatory. Alkali sulphates are known to react with silica at temperatures well above  $1000^{\circ}\text{C}.$ (7, 8) and Fialkov(8) reports that the reaction is catalysed by ferric oxide but not appreciably by alumina. The check experiment (Section II) shows that at  $875^{\circ}\text{C}.$  the reaction cannot be detected by X-ray diffraction.

The answer to the low temperature formation of sillimanite and orthoclase is doubtless to be sought in the form in which the silica is present in Chandler alunite. This is at present unknown. X-ray and electron diffraction methods have failed to reveal any colloidal silica or enough quartz, and no clue has been provided by the electron microscope. It has been suggested\* that the small alunite crystals are coated with a thin film of colloidal silica; but the electron diffraction evidence is negative, and the electron micrographs do not suggest any lack of sharpness of the edges of the alunite crystals. The only remaining possibility seems to be that the silica forms part of the alunite lattice itself. It is not known whether  $\text{Al}^{+++}$  ions in the alunite lattice at positions of coordination number 4 or 6 can be replaced isomorphously by  $\text{Si}^{++++}$ (9). In order to maintain the electrostatic charge balance, such a replacement would probably require the simultaneous replacement of an equal number of  $\text{Al}^{+++}$  sites of coordination number 6 by  $\text{Mg}^{++\dagger}$  or some similar ion. Whether isomorphous replacements of this kind can occur with no apparent effect on the X-ray diffraction pattern is a problem beyond the scope of this paper.

## VI. ACKNOWLEDGMENTS

The authors wish to acknowledge the interest and support of the State Alunite Industry, and the cooperation of the Physics Department, University of Western Australia, and of the Division of Industrial Chemistry, C.S.I.R. A grant from the Commonwealth Research Fund to one of us (N.S.B.) helped to defray the cost of the investigation.

## VII. REFERENCES

- (1) ASADA, Y.—*Sci. Pap. Inst. Phys. Chem. Res. Tokyo* **34**: 1201 (1938).
- (2) FINK, W. L., VAN HORN, K. R., and PAZOUR, H. A.—*Industr. Engng. Chem.* **23**: 1248 (1931).
- (3) BAYLISS, N. S., PAYNE, G. H., and PICKERING, R. W.—*Aust. Chem. Inst. J.* **10**: 38 (1943).

\* Dr. A. L. G. Rees, personal communication.

† Mr. F. Feakes, Chief Chemist of the State Alunite Industry, has written that some of the  $\text{Mg}^{++}$  in the ore is definitely present in an insoluble form.

- (4) BOWLEY, H.—Ann. Reps. Chemical Branch, Mines Dep., W. Aust., 1926, 1931.
- (5) STILLWELL, F. L.—Coun. Sci. Industr. Res. Aust., Mineragraphic Rep. No. 207 (March 1941).  
(This mimeographed report is partly quoted in reference (3) above.)
- (6) MELLOR, J. W.—“Comprehensive Treatise on Inorganic and Theoretical Chemistry”, Vol. 5, p. 353. (Longmans, Green & Co.: London, 1924.)
- (7) MELLOR, J. W.—Ibid. Vol. 6, p. 284.
- (8) FIALKOV, Y. A., and SHARGARODSKII, S. D.—*Mem. Inst. Chem. Ukrain. Acad. Sci.* **2**: 269 (1935). (*Chem. Abstr.* **31**: 4572 (1937).)
- (9) PAULING, L.—“The Nature of the Chemical Bond”, p. 382. (Cornell Univ. Press, 1940.)

## EXPLANATION OF PLATES 1-4

## PLATE 1

- Fig. 1.—Electron micrograph of alunite from Chandler, W.A., showing diatom fragment.  
Fig. 2.—Alunite from Chandler, W.A.

## PLATE 2

- Fig. 1.—Alunite dehydrated at 475° C. for 1 hour.  
Fig. 2.—Alunite dehydrated at 600° C. for 45 minutes.

## PLATE 3

- Fig. 1.—Alunite calcined under conditions for optimum yield of water-soluble potash, and then leached.  
Fig. 2.—Alunite “over-calcined” and not leached. Small needle-like crystals are possibly orthoclase.

## PLATE 4

- Alunite “over-calcined” and leached. Shows needle-like crystals tentatively identified as orthoclase.







# PLANE WAVES IN AN IONIZED GAS WITH STATIC ELECTRIC AND MAGNETIC FIELDS PRESENT

By V. A. BAILEY\*

[Manuscript received October 11, 1948]

## Summary

General equations are derived which specify such waves in a first approximation when their amplitudes are small. The equation of condition (19) (i.e. the equation of dispersion) is deduced and the nature of its roots is discussed. It is shown that under suitable conditions some of the roots lead to wave-groups which grow as they progress and others lead to waves which everywhere grow in time. The theory includes as special cases many of the results previously obtained for radio-propagation, "plasma" oscillations, and the like. Applications are indicated to electric oscillations and noise in discharge tubes, and to cosmic noise.

## I. INTRODUCTION

In the general theories which have been proposed to account for the observed properties of waves in ionized gases the effect of any static electric field which may be present in the medium has not been considered until recently(1). It is proposed here to develop such a theory which assumes that a static electric field is present and consequently takes account of the velocities of drift and the temperatures of the electrons and positive ions in the gas(2).

This theory can also be applied to the observations which have been made over many years(3, 4, 5) that under certain conditions internally controlled electric oscillations occur in low pressure discharge tubes and to the more recent observations of irregular electric waves in such tubes(6, 7, 8, 9, 10) and similar waves emitted by the ionosphere(11), the sun, and certain regions of our galaxy(12, 13).

More specifically the subject of our study is concerned with the wave properties of a medium which, in the steady state, consists of  $N_0$  electrons per cc. of temperature  $\theta$  and mean drift velocity  $U_0$ ,  $N_{i0}$  positive ions per cc. of temperature  $\theta_i$  and drift velocity  $U_{i0}$  and of gas molecules, and is pervaded by a static electric field  $E_0$  and a static magnetic field  $H_0$ .

The quantities  $\theta$ ,  $U_0$ ,  $\theta_i$ ,  $U_{i0}$  are functions of  $E_0$ ,  $H_0$  and the kind of gas and positive ions concerned.

## II. THE FUNDAMENTAL EQUATIONS

The basis of our theory consists of the following laws of physics :

- (1) Maxwell's laws of the electromagnetic field.
- (2) The conservation of electrons and positive ions.

\* Physics Department, University of Sydney.

(3) Maxwell's laws of the transfer of momentum in mixtures of different kinds of particles.

These laws provide the system (S) of five vector and three scalar equations, given below under (1) to (8), which relate the two particle-densities,  $N, N_i$ , the two particle-velocities  $\mathbf{U}, \mathbf{U}_i$ , the electric and magnetic vectors  $\mathbf{E}, \mathbf{H}$ , the scalar potential  $V$ , and the vector potential  $\mathbf{A}$ .

The values of all these quantities when the medium is in the steady state will be indicated by means of the suffix o. Their perturbations will be denoted by the corresponding lower-case letters; thus

$$N = N_0 + n, \quad N_i = N_{i0} + n_i, \quad \mathbf{U} = \mathbf{U}_0 + \mathbf{u}, \dots, \quad \mathbf{A} = \mathbf{A}_0 + \mathbf{a} \dots \dots \dots (\text{A})$$

The temperatures  $\theta, \theta_i$  will here be treated as constants.\*

The fundamental equations (S) are as follows:

$$\mathbf{E} = -c^{-1}D_t\mathbf{A} - \nabla V, \dots \dots \dots (1)$$

$$\mathbf{H} = \nabla \times \mathbf{A}, \dots \dots \dots (2)$$

$$4\pi e(N - N_i) = -\square^2 V, \dots \dots \dots (3)$$

$$4\pi e(N\mathbf{U} - N_i\mathbf{U}_i) = -c\square^2\mathbf{A}, \dots \dots \dots (4)$$

$$D_t N = -\nabla(N\mathbf{U}), \dots \dots \dots (5)$$

$$D_t N_i = -\nabla(N_i\mathbf{U}_i), \dots \dots \dots (6)$$

$$D\mathbf{U} + \nu\mathbf{U} + N^{-1}\nabla(\tau N) = (e/m)(\mathbf{E} + c^{-1}\mathbf{U} \times \mathbf{H}), \dots \dots \dots (7)$$

$$D\mathbf{U}_i + \nu_i\mathbf{U}_i + N_i^{-1}\nabla(\tau_i N_i) = -(e/m_i)(\mathbf{E} + c^{-1}\mathbf{U}_i \times \mathbf{H}), \dots \dots \dots (8)$$

where  $e$  = the charge on an electron =  $-4.80 \times 10^{-10}$  e.s.u.;

$m, m_i$  = masses of an electron and ion respectively;

$$\tau = \frac{1}{3}u_a^2 = k\theta/m = 1.515 \times 10^{11}\theta,$$

$\tau_i = \frac{1}{3}u_{ia}^2 - k\theta_i/m_i$ , where  $u_a^2, u_{ia}^2$  are the mean square velocities of agitation of the electrons and ions respectively;

$\nu, \nu_i$  = frequencies of collision, with gas molecules, of an electron and ion respectively;

$$D_t = \partial/\partial t, \quad D = D_t + (\mathbf{U}\nabla), \quad D_i = D_t + (\mathbf{U}_i\nabla);$$

$\nabla = \mathbf{i}\partial/\partial x + \mathbf{j}\partial/\partial y + \mathbf{k}\partial/\partial z$ , where  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  are unit orthogonal vectors parallel to the  $x, y$ , and  $z$  axes respectively;

$$\square^2 = \nabla^2 - c^{-2}D_t^2,$$

$c$  = velocity of light in vacuo.

A useful alternative to either (5) or (6) is the equation

$$D_t V = -c\nabla \cdot \mathbf{A} \dots \dots \dots (9)$$

In equations (7) and (8) the effects of collisions between electrons and positive ions are neglected. They need to be included only when  $N_0$  and  $N_{i0}$  are very large.

The system of equations ( $S_0$ ) for the steady state of the medium is immediately obtained from the system (S) by adding the suffix o to all the variables and omitting all the terms in which  $D_t$  occurs. When the expressions

\* To take account of their perturbations would require the addition of two more equations. This further complication of the theory is not justified at this stage for it would be analogous to replacing Newton's theory of sound waves by Laplace's which gives a value for the velocity of sound in air (for example) only about 10 per cent. greater than that given by Newton's theory.

in (A) for  $N$ ,  $N_i$ ,  $\mathbf{U}$ , etc. are substituted in the system (S) and the corresponding equations in the system ( $S_0$ ) are subtracted, we obtain the system of equations ( $S_p$ ) for the perturbations. As some of these are non-linear we may in a first approximation, or for sufficiently small perturbations, retain only the linear terms.

### III. THE LINEAR EQUATIONS FOR PLANE PERTURBATIONS

We will now consider the plane perturbations which can occur when the medium and the static fields are uniform in the steady state, i.e. when the quantities with a subscript 0 are all constants and  $N_{i0}=N_0$ . Accordingly we take all the lower-case variables to be proportional to

$$e^{i(\omega t - lx)} \dots\dots\dots (10)$$

The linear system of equations ( $S_p$ ) then reduces to the following system :

$$\mathbf{e} = -ic^{-1}\omega\mathbf{a} + ilv\mathbf{i}, \dots\dots\dots (1.1)$$

$$\mathbf{h} = -il\mathbf{i} \times \mathbf{a}_T, \dots\dots\dots (2.1)$$

$$n - n_i = (Z/4\pi ec^2)v, \dots\dots\dots (3.1)$$

$$N_0\mathbf{u} + \mathbf{U}_0n - N_0\mathbf{u}_i - \mathbf{U}_{i0}n_i = (Z/4\pi ec)\mathbf{a}, \dots\dots\dots (4.1)$$

$$Rn = N_0lu_1, \dots\dots\dots (5.1)$$

$$R_in_i = N_0lu_{i1}, \dots\dots\dots (6.1)$$

$$\delta\mathbf{u} - \Omega_0 \times \mathbf{u} - i\tau N_0^{-1}ln\mathbf{i} = (e/m)(\mathbf{e} + c^{-1}\mathbf{U}_0 \times \mathbf{h}), \dots\dots\dots (7.1)$$

$$\delta_i\mathbf{u}_i - \Omega_{i0} \times \mathbf{u}_i - i\tau_i N_0^{-1}ln_i\mathbf{i} = (-e/m_i)(\mathbf{e} + c^{-1}\mathbf{U}_{i0} \times \mathbf{h}), \dots\dots\dots (8.1)$$

$$\omega v = cla_1, \dots\dots\dots (9.1)$$

$$\left. \begin{array}{l} \text{where } Z = c^2l^2 - \omega^2, \quad R = \omega - U_1l, \quad R_i = \omega - U_{i1}l, \\ \delta = v + iR, \quad \delta_i = v_i + iR_i, \\ \Omega_0 = (-e/mc)\mathbf{H}_0, \quad \Omega_{i0} = (e/m_i c)\mathbf{H}_0, \end{array} \right\} \dots\dots\dots (11)$$

and the suffixes 1,  $T$  respectively denote components along and transverse to  $Ox$ .

On eliminating  $\mathbf{e}$ ,  $\mathbf{h}$ ,  $n$ ,  $n_i$  from (7.1) by means of the equations (1.1) to (6.1) and (9.1) we obtain the following equation in  $\mathbf{u}$  and  $\mathbf{u}_i$  alone.

$$\begin{aligned} & -iPZ\{(\delta - i\tau l^2 R^{-1})u_1 + \Omega_3 u_2 - \Omega_2 u_3\}\mathbf{i} \\ & -iPZ(-\Omega_3 u_1 + \delta u_2 + \Omega_1 u_3)\mathbf{j} \\ & -iPZ(\Omega_2 u_1 - \Omega_1 u_2 + \delta u_3)\mathbf{k} \\ & + (aR^{-1}u_1 + lU_2 u_2 + lU_3 u_3)\mathbf{i} \\ & + (lU_2 u_1 + Ru_2)\mathbf{j} \\ & + (lU_3 u_1 + Ru_3)\mathbf{k} \\ & - (bR_i^{-1}u_{i1} + lU_2 u_{i2} + lU_3 u_{i3})\mathbf{i} \\ & - (RR_i^{-1}lU_{i2} u_{i1} + Ru_{i2})\mathbf{j} \\ & - (RR_i^{-1}lU_{i3} u_{i1} + Ru_{i3})\mathbf{k} = 0. \dots\dots\dots (12) \end{aligned}$$

where  $P = m/4\pi e^2 N_0$ ,  $_{\tau}a = l^2 U_T^2 - Z$ ,  $b = l^2 U_T$   $U_{iT} - Z$ .

On equating to 0 the coefficients of  $\mathbf{i}$ ,  $\mathbf{j}$ ,  $\mathbf{k}$  and setting

$$u_1 = Rs_1, \quad u_{i1} = R_i s_{i1}, \dots\dots\dots (13)$$

we obtain

$$\left. \begin{array}{l} \alpha_1 s_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_{i1} s_{i1} + \alpha_{i2} u_{i2} + \alpha_{i3} u_{i3} = 0, \\ \beta_1 s_1 + \beta_2 u_2 + \beta_3 u_3 + \beta_{i1} s_{i1} + \beta_{i2} u_{i2} + \beta_{i3} u_{i3} = 0, \\ \gamma_1 s_1 + \gamma_2 u_2 + \gamma_3 u_3 + \gamma_{i1} s_{i1} + \gamma_{i2} u_{i2} + \gamma_{i3} u_{i3} = 0, \end{array} \right\} \dots\dots\dots (14)$$

where the coefficients  $\alpha_1$ ,  $\alpha_2$ ,  $\dots$   $\gamma_{i2}$ ,  $\gamma_{i3}$  are defined below under (16) and (17).

Similarly from (8.1) we obtain

$$\left. \begin{aligned} A_1 s_1 + A_2 u_2 + A_3 u_3 + A_{i1} s_{i1} + A_{i2} u_{i2} + A_{i3} u_{i3} &= 0, \\ B_1 s_1 + B_2 u_2 + B_3 u_3 + B_{i1} s_{i1} + B_{i2} u_{i2} + B_{i3} u_{i3} &= 0, \\ C_1 s_1 + C_2 u_2 + C_3 u_3 + C_{i1} s_{i1} + C_{i2} u_{i2} + C_{i3} u_{i3} &= 0, \end{aligned} \right\} \dots \dots \dots (15)$$

where  $A_1, A_2, \dots, C_{i2}, C_{i3}$  are defined below.

The coefficients of  $s_1, u_2, u_3$  in (14) and (15) are collected together below under (16) and those of  $s_{i1}, u_{i2}, u_{i3}$  are collected under (17).

$$\left. \begin{aligned} \alpha_1 &= -PZ(i\delta R + \tau l^2) + a, & \alpha_2 &= lU_2 - iZ\varphi_3, & \alpha_3 &= lU_3 + iZ\varphi_2, \\ \beta_1 &= R(lU_2 + iZ\varphi_3), & \beta_2 &= -iPZ\delta + R, & \beta_3 &= -iZ\varphi_1, \\ \gamma_1 &= R(lU_3 - iZ\varphi_2), & \gamma_2 &= iZ\varphi_1, & \gamma_3 &= -iPZ\delta + R, \\ A_1 &= -b, & A_2 &= -lU_{i2}, & A_3 &= -lU_{i3}, \\ B_1 &= -RilU_2, & B_2 &= -R_i, & B_3 &= 0, \\ C_1 &= -RilU_3, & C_2 &= 0, & C_3 &= -R_i, \end{aligned} \right\} (16)$$

$$\left. \begin{aligned} \alpha_{i1} &= -b, & \alpha_{i2} &= -lU_{i2}, & \alpha_{i3} &= -lU_{i3}, \\ \beta_{i1} &= -RlU_{i2}, & \beta_{i2} &= -R, & \beta_{i3} &= 0, \\ \gamma_{i1} &= -RlU_{i3}, & \gamma_{i2} &= 0, & \gamma_{i3} &= -R, \\ A_{i1} &= -P_i Z(i\delta_i R_i + \tau_i l^2) + a_i, & A_{i2} &= -lU_{i2} + iZ\varphi_3, & A_{i3} &= lU_{i3} - iZ\varphi_2, \\ B_{i1} &= R_i(lU_{i2} - iZ\varphi_3), & B_{i2} &= -iP_i Z\delta_i + R_i, & B_{i3} &= iZ\varphi_1, \\ C_{i1} &= R_i(lU_{i3} + iZ\varphi_2), & C_{i2} &= -iZ\varphi_1, & C_{i3} &= -iP_i Z\delta_i + R_i, \end{aligned} \right\} \dots \dots \dots (17)$$

$$\left. \begin{aligned} \text{where } P &= m/4\pi e^2 N_0, & P_i &= m_i/4\pi e^2 N_0, \\ a &= l^2 U_T^2 - Z, & b &= l^2 \mathbf{U}_T \mathbf{U}_T - Z, & a_i &= l^2 U_{iT}^2 - Z, \\ \varphi &= (\varphi_1, \varphi_2, \varphi_3) = -\mathbf{H}_0/4\pi e c N_0. \end{aligned} \right\} \dots \dots \dots (18)$$

As the six simultaneous equations (14) and (15) are linear and homogeneous the necessary and sufficient condition for their having a non-zero solution is that

$$\Delta \equiv \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_{i1} & \alpha_{i2} & \alpha_{i3} \\ \beta_1 & \beta_2 & \beta_3 & \beta_{i1} & \beta_{i2} & \beta_{i3} \\ \gamma_1 & \gamma_2 & \gamma_3 & \gamma_{i1} & \gamma_{i2} & \gamma_{i3} \\ A_1 & A_2 & A_3 & A_{i1} & A_{i2} & A_{i3} \\ B_1 & B_2 & B_3 & B_{i1} & B_{i2} & B_{i3} \\ C_1 & C_2 & C_3 & C_{i1} & C_{i2} & C_{i3} \end{vmatrix} \times Z^{-4} = 0 \dots \dots (19)$$

The determinant  $|\alpha_1, \beta_2, \dots, C_{i3}|$  in (19) is divided by  $Z^4$  in order to remove the factor  $Z^4$  which it always contains and which leads to waves that are entirely independent of the given medium. The factor  $Z^4$  appears immediately on transforming the determinant by means of the following operations:

- From col. 1 subtract  $lU_2 \times \text{col. 2}$  and  $lU_3 \times \text{col. 3}$ ;
- from col. 4 subtract  $lU_{i2} \times \text{col. 5}$  and  $lU_{i3} \times \text{col. 6}$ ;
- to col. 2 add col. 5 and to col. 3 add col. 6;
- from  $R \times \text{row 1}$  subtract  $lU_2 \times \text{row 2}$  and  $lU_3 \times \text{row 3}$ ;
- from  $R_i \times \text{row 4}$  subtract  $lU_{i2} \times \text{row 5}$  and  $lU_{i3} \times \text{row 6}$ .

The resulting determinant divided by  $RR_i$  and equated to zero yields an interesting alternative form of the equation of condition (19).

The quantity  $\Delta$  is a polynomial  $P(\omega, l)$  of degree 12 in either  $\omega$  or  $l$ .

#### IV. SOME ASPECTS OF THE EQUATION OF CONDITION

Equation (19) includes as very special cases the dispersion equation of the Magneto-Ionic theory of radio-propagation (14) and certain relations published by



Thomson(15), Tonks and Langmuir(5), Linder(16) and others, and so may be applied to the same fields as these.

When the angular frequency  $\omega$  has a given real value the 12 roots  $l_n$  of (19) specify the corresponding 12 different refractive indices and associated coefficients of attenuation of the medium. To study the nature of these roots it is usually necessary to proceed by approximations and also initially to consider certain standard cases. In the first approximation we may take  $\nu$  and  $\nu_i$  to be negligibly small while at the same time keeping  $U_0$  and  $U_{i0}$  unaltered\*; this reduces  $\Delta$  to  $\Delta_0$  a polynomial with all its coefficients real. Thus we take the roots  $l_{0n}$  of the equation

$$\Delta_0 = 0 \dots\dots\dots (19.0)$$

as first approximations to the roots  $l_n$ .

The following results have been found in the cases so far studied.

Most of the roots  $l_{0n}$  are real and, after taking the second approximations, the corresponding waves and wave-groups are of the familiar types which attenuate as they progress. But when  $\omega$  lies within certain finite frequency bands one or more pairs of roots are conjugate complex numbers like  $l_{01}$ ,  $l_{02} = \alpha \pm i\beta$ ; when  $\alpha \neq 0$  the important consequence follows that a narrow wave-group centred about  $\omega$  will grow in amplitude as it travels, and ultimately do so proportionately to  $e^{|\beta x|}$ .

On making the second approximations it is found that such growing wave-groups still occur when the collision frequencies  $\nu$ ,  $\nu_i$  are smaller than particular quantities. In limited circumstances these can also occur when  $\alpha = 0$ .

Similarly when the wave number  $l$  has a given real value the 12 roots  $\omega_n$  of (19) specify the frequencies and associated damping coefficients of the 12 possible corresponding waves. As first approximations to the roots  $\omega_n$  we may take the roots  $\omega_{0n}$  of (19.0), now regarded as an equation in  $\omega$ .

In the cases so far studied most of the roots  $\omega_{0n}$  are real and the corresponding waves are of the familiar type which damps out with the lapse of time. But when  $l$  lies within certain finite wave-number ranges at least one pair of roots are conjugate complex numbers like  $\omega_{01}$ ,  $\omega_{02} = \gamma \pm i\delta$  and then one of the corresponding waves everywhere grows in amplitude with the lapse of time proportionately to  $e^{|\delta t|}$ . These growing waves also occur after making the second approximations, provided that  $\nu$ ,  $\nu_i$  do not exceed certain quantities.

Waves which correspond to real roots,  $l_{0n}$  or  $\omega_{0n}$ , of (19.0) may conveniently be called those of the First Species and waves which correspond to conjugate complex roots may be called those of the Second Species. The first are of familiar type but the second include wave-groups growing in space, and waves growing in time. Examples of the second species are given below. These and other physical aspects of roots of the equation of condition will be considered more fully in another publication.

\*If this *mathematical* procedure should need clarifying by means of a corresponding physical process we may suppose that the static electric force  $E_0$  is simultaneously reduced to a suitable low value.

## V. APPROXIMATION WHEN THE MOTION OF THE POSITIVE IONS IS NEGLECTED

When the vibratory motion of the positive ions is negligible by comparison with that of the electrons (e.g. when  $\omega$  is large enough) the equation of condition (19) reduces to

$$\Delta_e \equiv \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} \times Z^{-1} = 0 \dots\dots\dots (20)$$

and is of the eighth degree in  $\omega$  and  $l$ .

On expanding  $\Delta_e$  and for convenience taking  $c$  as the unit of velocity, the electron-density frequency  $p$  as the unit of frequency and  $p/c$  as the unit of wave-number, (20) becomes

$$(X + U_T^2 l^2 Z^{-1})(Y^2 - \Omega_1^2 Z^2) - RY(\Omega_T^2 Z + U_T^2 l^2 Z^{-1}) - 2lRZ\Omega_1(\Omega_T \cdot U_T) = 0 \dots\dots\dots (21)$$

$$\left. \begin{aligned} \text{where } X &= R^2 - \tau l^2 - 1 - i\nu R, \\ Y &= ZR + R - i\nu Z, \\ \Omega_0 &= -H_0 e/mc. \end{aligned} \right\} \dots\dots\dots (22)$$

Equation (21) was derived more directly in an earlier publication(2).\*

In the standard case where

$$U_T = 0, \quad H_T = 0 \dots\dots\dots (C11)$$

i.e. where the electron drift-velocity, the magnetic field, and the direction of propagation are all parallel, (21) reduces to the three equations

$$R^2 - \tau l^2 - 1 - i\nu R = 0, \quad (\text{roots } l_1, l_2), \dots\dots\dots (23)$$

$$ZR + R - i\nu Z + \Omega_0 Z = 0, \quad (\text{roots } l_3, l_4, l_5), \dots\dots\dots (24)$$

$$ZR + R - i\nu Z - \Omega_0 Z = 0, \quad (\text{roots } l_6, l_7, l_8), \dots\dots\dots (25)$$

Table 1 gives numerical examples of roots  $l_n$  which yield growing wave-groups.

TABLE 1

(C11)

$p=5.65 \times 10^4 \sqrt{N_0}$ .

$\theta$ (°K.)	$U_0$ (cm./sec.)	$\Omega_0$	$\omega$	$\nu$	$l_n$ (in units equal to $p/c$ .)
		(in units equal to $p$ )			
$5.94 \times 10^6$	$3 \times 10^8$	—	0.01	0.1	$\begin{cases} l_1 = -1.11 - 32.77i \\ l_2 = -1.11 + 33.89i \end{cases}$
$1.2 \times 10^6$	$3 \times 10^8$	—	0.5	0.1	$\begin{cases} l_1 = -43 - 66i, \\ l_2 = -57 + 76i. \end{cases}$
—	$3 \times 10^9$	2.375	0.00125	neg. <sup>1e</sup>	$l_3, l_4 = 10^{-2}(2.45 \pm 0.525i)$
—	$3 \times 10^9$	1.732	1.26	neg. <sup>1e</sup>	$l_6, l_7 = -3.25 \pm 0.78i.$

In the standard case where

$$U_T = 0, \quad H_1 = 0 \dots\dots\dots (C12)$$

i.e. where the drift-velocity and direction of propagation are parallel and the magnetic field is perpendicular to both, (21) reduces to the equations

\* Submitted to Royal Society of N.S.W. on May 10, 1948.

$$XY - \Omega_0^2 RZ = 0, \text{ (roots } l_1, l_2, \dots, l_5) \dots\dots\dots (26)$$

$$Y = 0, \text{ (roots } l_6, l_7, l_8) \dots\dots\dots (27)$$

Both equations yield waves of the second species but the waves corresponding to (26) require conditions for growth which are much less restricted than those which correspond to (27). Numerical examples of roots which yield growing wave-groups when  $\tau$  and  $\nu$  are negligible and  $\Omega_0 = 1.732$ , are:

$$U_0 = 3 \times 10^8, \quad \omega = 1.95 : l_1, l_2 = -4.22 \pm 3i.$$

$$U_0 = 3 \times 10^9, \quad \omega = 1.62 : l_1, l_2 = -2.8 \pm 0.44i.$$

In general, growing wave-groups can occur in the cases (C11) and (C12) only when the frequency  $\omega$  is less than  $\frac{1}{2}(\Omega_0 + \sqrt{\Omega_0^2 + 4})$ .

## VI. MOTION OF THE POSITIVE IONS INCLUDED

When the vibrations of the positive ions are not neglected the calculations are more difficult, so only one example will be considered.

In the standard case where

$$U_T = 0, \quad U_{iT} = 0, \quad H_T = 0, \dots\dots\dots (C11)$$

the determinant  $\Delta$  in (19) has  $(\alpha_1 A_{i1} - Z^2)Z^{-4}$  as one of its factors and so four of the possible waves correspond to the roots of the equation

$$(R^2 - \tau l^2 - i\nu R - p^2)(R_i^2 - \tau_i l^2 - i\nu_i R_i - p_i^2) - p^2 p_i^2 = 0 \dots\dots\dots (28)$$

where  $p = \sqrt{4\pi N_0 e^2 / m}$  and  $p_i = \sqrt{4\pi N_0 e^2 / m_i}$ ;

$p$  and  $p_i$  may be termed respectively the electron-density frequency and the ion-density frequency.

When  $\nu$  and  $\nu_i$  are small, first approximations to the roots of (28) are given by the roots of

$$(R^2 - \tau l^2)(R_i^2 - \tau_i l^2) - p_i^2(R^2 - \tau l^2) - p^2(R_i^2 - \tau_i l^2) = 0 \dots\dots\dots (28.0)$$

If  $\omega$  and  $l$  are taken as the coordinates of a point, the curve of the fourth degree defined by (28.0) is easily drawn when  $U_1$ ,  $U_{i1}$ ,  $\tau$ ,  $\tau_i$ ,  $p$ , and  $p_i$  have assigned numerical values and may then be used to determine the frequency ranges within which conjugate complex roots  $l_{01}$ ,  $l_{02}$  occur and the wave-number ranges within which similar roots  $\omega_{01}$ ,  $\omega_{02}$  occur. The following is an illustrative example\* which has been worked out by Mr. J. A. Roberts at my suggestion.

When

$$\begin{aligned} p &= 10^8 \text{ rad./sec.}, & p_i &= 10^6 \text{ rad./sec.}, \\ U_1 &= 3 \times 10^8 \text{ cm./sec.}, & U_{i1} &= -3 \times 10^5 \text{ cm./sec.}, \\ \tau &= 10^{16}, & \tau_i &= 10^{12}, & m_i &= 10^4 m, \end{aligned}$$

we have the following results:

In the frequency ranges  $0-10^6$  and  $10^6-2 \times 10^6$  rad./sec. (approximately) we have respectively two pairs and one pair of conjugate complex roots  $l_{0n}$ ; in the wave-number range  $0-0.38$  rad./cm. we have one pair of conjugate complex roots  $\omega_{0n}$ . In particular when  $\omega = 1.16 \times 10^6$  then  $l_{01}$ ,  $l_{02} = -0.43 \pm 0.32i$  and when

$$l = 0.15 \text{ then } \omega_{01}, \omega_{02} = 10^4(-3.8 \pm 44.3i) \dots\dots\dots (29)$$

\* This and some of the other numerical examples are selected primarily to illustrate the mathematical processes and do not necessarily correspond exactly with known physical conditions.

The following example of solutions of (28) has been derived by means of a purely analytical method of approximation:

We take the following values which correspond very roughly to those in discharge in hydrogen

$$\begin{aligned} p &= 6.1 \times 10^8 \text{ rad./sec.}, & p_i &= 10^7 \text{ rad./sec.}, \\ v &= 10^8 \text{ coll./sec.}, & v_i &= 2 \times 10^6 \text{ coll./sec.}, \\ U_1 &= 1.5 \times 10^8 \text{ cm./sec.}, & U_{i1} &= -2 \times 10^6 \text{ cm./sec.}, \\ \tau &= 10^{16} \text{ cm.}^2/\text{sec.}^2, & \tau_i &= 4 \times 10^{10} \text{ cm.}^2/\text{sec.}^2, \end{aligned}$$

When  $l=0.2$  it is found that (28) has the following two roots:

$$\omega_1 = 10^6(-7.48 + 21.46i), \quad \omega_2 = 10^6(-0.357 - 1.203i) \dots\dots\dots (30)$$

The other factors of  $\Delta$  in this case lead to the two quartic equations

$$Z(-iP\delta \pm \varphi_1)(-iP_i\delta_i \mp \varphi_1) + (-iP\delta \pm \varphi_i)R_i + (-iP_i\delta_i \mp \varphi_1)R = 0 \dots (31)$$

The corresponding eight waves and other physical aspects of the theory given in this and other sections will be discussed in another publication.

## VII. APPLICATIONS OF THE THEORY

The general theory can be applied to all situations in which the postulated conditions actually exist. These include low-pressure discharge tubes, the ionosphere, the solar atmosphere and probably also other stellar atmospheres. In such situations when the conditions favour the occurrence of growing wave-groups almost any initial random disturbance will lead to much larger disturbances. As the numerical examples show, a strong magnetic field can extend the range of frequencies within which such growth occurs.

An essential condition for growth is that at least one drift velocity exists and this requires the existence of an accelerating electric field. Such a field can be provided either by means of a changing magnetic field, by motion of the ionized medium relative to a magnetic field or by the static charges on the electrodes and walls of a discharge tube. The static electric field may be regarded as supplying a source of energy for the generation of growing waves in the same way as a battery of cells provides energy for the generation of oscillations in circuits connected with electron tubes.

The resulting random growing wave-groups will have the general character of electric "noise" and will range in frequency from very low values to an upper limit. The observed "noise" from extra-terrestrial sources and from the ionosphere may be supposed to originate in part from such growing waves. For solar noise this is borne out by the fact that, according to Pawsey,\* the angular frequencies of the fluctuating, very intense components rarely exceed an upper limit of about  $3 \times 10^9$  rad./sec. Short-period terrestrial magnetic variations may in part have a similar origin(17).

When complex roots like those in (29) and (30) occur under conditions which define a particular length, such as those in a discharge tube with electrodes, it can be shown that certain initial disturbances will at every place increase with time. We may perhaps also then expect certain periodic oscillations in the tube to be maintained; but this last point needs further consideration.

\* Personal communication.



Detailed comparison of observations with the theory has to be made before reliable conclusions on these questions are reached.

Before concluding it should be mentioned that the general theory given here has some relation to the work on travelling-wave tube amplifiers by Pierce(18), Kompfner(19), Chu and Jackson(20), and Rydbeck(21). In these tubes the conditions are highly specialized, but in his discussion of the relevant theory, Pierce, in effect, makes use of pairs of conjugate complex roots of an equation to derive a spatially growing wave of a given frequency. More recently Pierce(22) has developed a similar theory to account for certain fluctuations of current in his tube but finds it necessary for this purpose to postulate the existence and motion of positive ions in his electron beam.

Important differences exist between his discussion of waves and ours. For example, he assumes that the mean velocity of the positive ions is in the *same* sense as the mean velocity of the electron beam. Also he does not consider the effects of the electron and ion temperatures. The theoretical formulae given here may, however, be regarded as at least *formally* including his results for plane waves (under our special case (C11)) when our quantity  $U_{i0}$  has its sign changed and the quantities  $E_0$ ,  $\tau$ ,  $\tau_i$ ,  $\nu$ , and  $\nu_i$  are all given the value zero.

#### VIII. ACKNOWLEDGMENT

In conclusion I desire to acknowledge the great benefits this work has received from the critical discussions and other assistance which have been most willingly given by Dr. R. E. B. Makinson and Messrs. P. G. Guest and J. A. Roberts.

#### IX. REFERENCES

- (1) BAILEY, V. A.—*Nature* **161**: 599 (1948).
- (2) BAILEY, V. A.—*J. Roy. Soc. N.S.W.* **82**: 107 (1948).
- (3) APPLETON, E. V., and WEST, A. G.—*Phil. Mag.* **45**: 879 (1923).
- (4) PENNING, F. M.—*Nature* **118**: 301 (1926).
- (5) TONKS, L., and LANGMUIR, I.—*Phys. Rev.* **33**: 195 (1929).
- (6) RUTHBERG, W. T.—*Ibid.* **70**: 112 (1946).
- (7) COBINE, J. D., and GALLAGHER, C. J.—*Ibid.* **70**: 113 (1946).
- (8) THONEMANN, P. C., and KING, R. B.—*Nature* **158**: 414 (1946).
- (9) DENISSE, J., and STEINBERG, J. L.—*C.R. Acad. Sci. Paris* **224**: 646 (1947).
- (10) STEINBERG, J. L.—*Nature* **160**: 833 (1947). *Rev. Sci. Paris* **85**: 601 (1947).
- (11) COVINGTON, A. E.—*Terr. Magn. Atmos. Elect.* **52**: 339 (1947).
- (12) REBER, G., and GREENSTEIN, J. L.—*Observatory* **67**: 836, 15 (1947).\*
- (13) DENISSE, J.—*Rev. Sci. Paris* **85**: 483 (1947).\*
- (14) MIMNO, H. R.—*Rev. Mod. Phys.* **9**: 19, 20 (1937).\*
- (15) THOMSON, J. J.—*Phil. Mag.* **6**: 1254 (1928).
- (16) LINDER, E. G.—*Phys. Rev.* **49**: 753 (1936).
- (17) WILLIS, H. F.—*Nature* **161**: 887 (1948).
- (18) PIERCE, J. R.—*Proc. Inst. Radio Engrs.* **35**: 111 (1947).
- (19) KOMPFFNER, R.—*Ibid.* **35**: 124 (1947).
- (20) CHU, I. J., and JACKSON, D.—Res. Lab. Electronics, Mass. Inst. Technol. Techn. Rep. No. 38 (1947).
- (21) RYDBECK, O. E.—Res. Lab. Electronics, Chalmers Inst. Technol. Rep. No. 1 (1947).
- (22) PIERCE, J. R.—*J. Appl. Phys.* **19**: 231 (1948).

\* These give summaries of the work done in their respective fields.



# HYDROGEN ATMOSPHERES IN THE ABSENCE OF THERMODYNAMIC EQUILIBRIUM

## IV. THE SOLAR CHROMOSPHERE

By R. G. GIOVANELLI\*

[Manuscript received August 27, 1948]

### Summary

A discussion of the observational data shows that in the range  $0 < z < 2.5 \times 10^8$  cm. the electron concentration in the lower chromosphere may be expressed as

$$N_e = 5 \times 10^{11} \exp(-6 \times 10^{-9} z) \text{ per cc.},$$

where  $z$  cm. is the height measured from a level 500 km. above the base of the chromosphere. It is shown that above the 500-km. level hydrogen atoms are almost completely ionized, and that with hydrostatic equilibrium the observed density gradient corresponds to a temperature of  $2.7 \times 10^4$  °K. This is in good agreement with the temperature derived by Redman from line profile measurement, so that the lower chromosphere appears to be effectively in thermal and hydrostatic equilibrium.

Estimates of lower accuracy are given for the electron concentrations at some higher levels in the chromosphere.

A discussion is given of the mechanism of absorption of H $\alpha$  radiation by a chromosphere whose temperature is well above that of the photosphere.

## I. INTRODUCTION

In recent years it has become increasingly evident that the solar chromosphere is a region hotter than the underlying photosphere. Many of the observations require reconsideration in light of this fact, and it is shown in this paper that a consistent picture of the electron concentration and temperature may be obtained from the observations of Menzel, Cillié and Menzel, Redman, and Mitchell.

## II. ABSOLUTE ELECTRON CONCENTRATIONS

The electron concentration may be estimated by two independent methods, which we shall consider in turn.

(a) Suppose the variation in electron concentration,  $N_e$ , with height to be given by

$$N_e = (N_e)_0 \exp(-\beta z) \dots\dots\dots (1)$$

where  $z$  is the height (in cm.) and  $\beta$  is the concentration gradient. The emission in the Balmer continuum is determined by the proton and electron concentrations (which are almost equal) and by the electron temperature  $\theta$ , and falls off with a gradient  $2\beta$  if  $\theta$  be uniform with height. Cillié and Menzel(1) found  $2\beta$  from

\* Division of Physics, C.S.I.R.

observations at elevations of 670 and 1500 km. above the base of the chromosphere. Assuming a constant gradient, they deduced from observations of the Balmer continuum at wavelength 3640 Å. that, at the base of the chromosphere,

$$\log N_+ N_e = 17.16 + (3/2) \log \theta \dots\dots\dots (2)$$

If  $\theta = 2.5 \times 10^4$  °K., as is later shown to be close to the temperature of the lower chromosphere,  $N_e = 7.6 \times 10^{11}$  per cc.

This result is obtained by using an extrapolation formula to deduce concentrations at the base of the chromosphere from observations at 670 km. It is not certain that such a large extrapolation is warranted near the base of the chromosphere. Using Cillié and Menzel's gradient,  $\beta = 0.77 \times 10^{-8}$  cm.<sup>-1</sup>, we find for  $\theta = 2.5 \times 10^4$  °K.

$$(N_e)_{670} = 4.5 \times 10^{11} \text{ per cc.}$$

For comparison with concentrations obtained below from other observations, we make a slight extrapolation to 500 km., and find

$$(N_e)_{500} = 5.2 \times 10^{11} \text{ per cc.}$$

(b) Inglis and Teller(2), in a rediscussion of Pannekoek's theory of the Stark broadening of Balmer lines by inter-ionic fields(3), showed that for a temperature greater than  $10^5/n_u$ , as in the chromosphere,

$$\log N_+ = 23.26 - 7.5 \log n_u \dots\dots\dots (3)$$

$n_u$  being the principal quantum number of the last resolved series line.

The major error arising from the use of this formula comes from uncertainty in deciding which is the last resolved line in the spectral series. Wildt(4), using Mitchell's recorded lines(5), found at a height of 500 km.  $N_+$  ( $=N_e$ ) to be  $3.16 \times 10^{11}$  or  $1.74 \times 10^{11}$  per cc., depending on whether Balmer or Paschen series observations be used; the geometric mean of these is  $2.25 \times 10^{11}$  per cc. From Cillié and Menzel's 1932 eclipse observations(1), in which the last resolved Balmer series line at 670 km. was  $n=31$ , we find  $N_e = 1.2 \times 10^{12}$  per cc. The last Balmer series line recorded with certainty, at a height of 300 km., by Menzel(6) was  $n=31$ , although doubtful lines were recorded to  $n=37$ . The corresponding values of  $N_e$  are  $1.2 \times 10^{12}$  and  $3.2 \times 10^{11}$  per cc. respectively, the geometric mean being  $6.2 \times 10^{11}$  per cc.

To adjust for observations made at slightly different heights, we may use the equation of hydrostatic equilibrium which, for  $\theta = 2.5 \times 10^4$  °K., and a chromosphere consisting mainly of ionized hydrogen (as we later find to be the case), becomes

$$N_e = N_0 \exp(-0.65 \times 10^{-8} z) \dots\dots\dots (4)$$

Values of  $N_e$  at the 500-km. level as obtained by the two methods are given in Table 1. The various observations and methods of reduction are in agreement to within the limits of observation error, the geometric mean being  $5.3 \times 10^{11}$  per cc. We thus adopt as the electron concentration at this level

$$N_e = 5 \times 10^{11} \text{ per cc.} \dots\dots\dots (5)$$

This value is not significantly affected by uncertainties in the adopted value of  $\theta$ , and is probably reliable to within a factor of less than 2.

TABLE 1  
ELECTRON CONCENTRATIONS AT 500 KM. ( $\theta = 2.5 \times 10^4$  °K.)

Observer	Method	$N_e$ (per cc.)
Cillié and Menzel ..	Photometry of Balmer continuum	$5.2 \times 10^{11}$
Cillié and Menzel ..	Last resolved line	$1.3 \times 10^{12}$
Menzel .. .. .	" " "	$5.4 \times 10^{11}$
Mitchell. . . . .	" " "	$2.2 \times 10^{11}$
Geometric mean ..	.. .. .	$5.3 \times 10^{11}$

### III. THE VARIATION OF ELECTRON CONCENTRATION WITH HEIGHT

The monochromatic Balmer emission of a 1-cm. thick vertical section of the chromosphere (assumed isothermal) above a plane tangent to the moon's limb is

$$Q = LA_{n,2}n^2N_0^2 \exp(1.58 \times 10^5/n^2\theta) \text{ quanta per sec.} \dots\dots\dots (6)$$

where the Balmer line originates from a high atomic level of principal quantum number  $n$ ;  $A_{n,2}$  is the Einstein coefficient of spontaneous transition;  $N_0$  is the electron concentration at the lowest visible level in the sun's atmosphere (Fig. 1); and  $L$  is a quantity independent of  $n$ ,  $N_0$ , and  $\theta$  (e.g. see (4)).

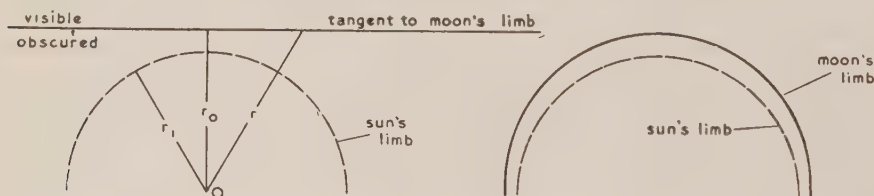


Fig. 1.—Region of sun's atmosphere visible during eclipse.

The above expression applies only to lines for which  $N_n$ , the number of hydrogen atoms in the  $n$ -quantum state, is related to the electron concentration as it would be in thermodynamic equilibrium at the electron temperature. Menzel and Cillié(7) found this condition to hold accurately in the chromosphere for  $n > 22$ , less accurately for  $n < 22$ . For an assumed  $\theta$  of  $2 \times 10^4$  °K., they found a deficiency of atoms in the 5-quantum state amounting to a factor of 2.6. The author has shown(8), from a discussion of transition rates, that the error decreases as  $n$  increases, being less than a factor of 2 for  $n > 10$ .

Using equation (6), the electron density gradient may be derived either from Cillié and Menzel's results or from Mitchell's observations.

(a) Cillié and Menzel(1) deduced from observations of the Balmer continuum emitted by the chromosphere above the two heights of 670 and 1500 km. that the electron density gradient  $\beta$  is  $0.77 \times 10^{-8}$  per cm.

It is also possible to deduce  $\beta$  from their measured Balmer line intensities. In Table 2 we have summarized the observed variations of  $Q$  with elevation for two groups of lines. In the lowest regions the gradient of  $Q$  differs for different

TABLE 2  
RATIOS OF MONOCHROMATIC EMISSIONS ABOVE VARIOUS CHROMOSPHERIC LEVELS

Lines	Mean $\frac{Q_{670}}{Q_{1500}}$	Lines	Mean $\frac{Q_{1500}}{Q_{2330}}$	Lines	Mean $\frac{Q_{2330}}{Q_{3170}}$
H <sub>10</sub> -H <sub>17</sub> .. ..	6.58	H <sub>10</sub> -H <sub>17</sub> .. ..	6.46	H <sub>10</sub> -H <sub>13</sub> ..	6.83
H <sub>20</sub> -H <sub>29</sub> .. ..	7.36				

Note.— $Q_z$  is the emission above a plane passing  $z$  km. above the base of the chromosphere.

values of  $n$ , perhaps due to slight departures from the predictions of the equations of thermodynamic equilibrium, or to self absorption. If observations on the lowest regions be used to obtain a correction factor for the ratios for  $10 < n < 17$ , then over the whole range 670-3170 km. a constant value  $\beta = 0.77 \times 10^{-8}$  per cm. is found. But if we rely without correction on lines for which  $10 < n < 17$ , then  $\beta$  is found to be  $0.63 \times 10^{-8}$  per cm.

(b) We pass to deductions from data of the heights to which hydrogen lines are observed(5). These have been carefully considered previously by Wildt(4), but small corrections are required because he assumed an excitation temperature of 4830° K.

Following Mitchell and Williams(9), and Wildt(4), we have assumed that the observed height of a spectral line corresponds to an energy  $E = Qh\nu$ , which is constant from line to line, associated with the number of quanta emitted per second,  $Q$  (equation (6)). We have also adopted Wildt's systematic wavelength corrections to this expression. Putting  $\theta = 3 \times 10^4$  °K., the value of  $\log N_0^2 + \text{constant}$  may be plotted against the observed height of the same line. The resulting curve is similar to Wildt's, and it reduces to a straight line for heights below 3000 km. ( $n \geq 19$ ), the corresponding value of  $\beta$  being  $0.425 \times 10^{-8}$  per cm. This gradient is a little more than one-half that derived in (a). There is other evidence that physical conditions in the chromosphere are by no means constant from eclipse to eclipse(1, 10), while the chromosphere appears far from uniform in spectroheliograms. Adopting the mean of the gradients calculated above, we find the electron concentration to be given by

$$N_e = 5 \times 10^{11} \exp\{-0.60 \times 10^{-8}z\} \text{ per cc.} \dots\dots\dots (7)$$

where  $z$  is the height, in cm., measured from a layer 500 km. above the base of the chromosphere.

Mitchell's height measurements may be used to derive very roughly, on the assumption of uniform temperature, the electron concentrations at various



higher elevations relative to that at 3000 km. (Table 3). Since the populations in the chromosphere of the states for  $n \approx 10$  cannot be calculated from the equation of thermodynamic equilibrium to within a factor of less than 2, we shall not use

TABLE 3  
ELECTRON CONCENTRATIONS AT ELEVATIONS ABOVE 3000 KM.

Height (km.)	$N_e/N_{3000}$	$N_e$ • (per cc.)
3500	0.84	$9.4 \times 10^{10}$
4000	0.73	$8.2 \times 10^{10}$
4500	0.69	$7.7 \times 10^{10}$
5000	0.65	$7.3 \times 10^{10}$
5500	0.60	$6.7 \times 10^{10}$
6000	0.44	$4.9 \times 10^{10}$

the results for  $n < 12$ . The ratios derived are likely to be in error by a factor which increases with height, but should not exceed  $\sqrt{2}$  at 6000 km. unless the temperature be non-uniform.

There is one other line from which a very rough estimate may be made, viz.  $H\alpha$ , which may be traced to 12,000 km. The emission of  $H\alpha$  quanta in the absence of thermodynamic equilibrium has recently been examined (11). Over the range  $2.5 \times 10^4 \leq \theta \leq 1.0 \times 10^5$  °K., the number of quanta emitted may be expressed approximately as  $4.6 \times 10^{-7} N_e^2 / \theta^{3/2}$  per cc. per sec. provided the spectral radiation densities do not exceed those of a black body at a temperature of 5000° K., whereas in thermodynamic equilibrium it is given by

$$4.2 \times 10^{-16} \times 3^2 N_e^2 \exp(1.76 \times 10^4 / \theta) \cdot A_{3,2} / \theta^{3/2} \text{ per cc. per sec.}$$

Thus for  $H\alpha$ , equation (6) may be replaced by

$$Q = L \cdot N_0^2 (4.6 \times 10^{-7} / 4.2 \times 10^{-16}) \dots \dots \dots (8)$$

At the highest level at which any line is observed,  $Q = E / h\nu C_\lambda$ , where  $E$  is constant from line to line and  $C_\lambda$  is the colour correction factor given by Wildt. But since we now have evaluated  $N_e$  for the lower chromosphere, the mean value of  $E$  may be deduced from the lines  $n > 12$ . In this way we find the value of  $N_e$  at 12,000 km. to be  $7.7 \times 10^9$  per cc.

The electron concentrations derived here for heights above about 3000-4000 km. are only rough estimates. Not only is the corresponding concentration of neutral atoms in the appropriate excited state used in the deductions not known to great accuracy, but equation (6) applies only for  $\theta$  constant. Moreover, it is by no means certain that the significance of measurements of the heights of spectral lines is yet fully understood.



## IV. ELECTRON CONCENTRATIONS IN THE CORONA

It is of interest to compare electron concentrations in the chromosphere and corona. According to Allen(12), who has revised Baumbach's coronal electron densities(13), the electron density at a distance  $r$  from the centre of the sun (expressed in units of the sun's radius) is

$$N_e = 10^8(1.55r^{-6} + 2.99r^{-16}) \dots\dots\dots (9)$$

The electron concentrations as given by this formula are tabulated in Table 4 for elevations from  $2.5 \times 10^4$  down to  $1.4 \times 10^4$  km. It is not known how close

TABLE 4

ELECTRON CONCENTRATIONS IN THE INNER CORONA  
(FROM ALLEN'S FORMULA)

Height above limb ( $10^4$ km.) ..	2.5	2.0	1.8	1.6	1.4
Electron concentration ( $10^8$ per cc.)	2.94	3.21	3.31	3.43	3.54

to the sun's surface this equation is valid, but it is fairly safe to assume it to hold at heights of 20,000 km. and higher. Comparison with results for the chromosphere shows that it cannot hold down to 12,000 km. and, moreover, that the temperature at 12,000 km. is still small compared with that of the corona.

## V. THE TEMPERATURE DISTRIBUTION AND STATE OF IONIZATION

The temperature of the chromosphere, if it be in hydrostatic equilibrium, may be derived from the electron densities. For an almost completely ionized atmosphere, the equation of hydrostatic equilibrium gives

$$N_e = (N_0 \theta_0 / \theta) \exp \left\{ -1.63 \times 10^{-4} \int_0^z \theta^{-1} dz \right\} \dots\dots\dots (10)$$

It seems very likely that throughout the region  $0 < z < 2500$  km., where the concentration gradient is constant, the temperature is also constant. With this assumption,  $\theta = 2.71 \times 10^4$  °K., in very good agreement with the temperature of  $3.0 \times 10^4$  °K. deduced by Redman(14) from profiles of spectral lines at an estimated altitude of 1500 km.

The state of ionization in the chromosphere above the 500-km. level is readily deduced; for it has been shown(15) that hydrogen at a temperature of  $2.7 \times 10^4$  °K. and a concentration of  $5 \times 10^{11}$  per cc. is highly ionized. A chromosphere of predominantly neutral hydrogen atoms and having the observed ion concentrations would need a temperature of  $5.4 \times 10^4$  °K. for its support(15); but hydrogen at this temperature and with these ion concentrations is almost completely ionized. Thus the chromosphere must be highly ionized if in hydrostatic equilibrium.

From a knowledge of the electron concentrations and temperatures at two elevations it is possible to derive from the equation of hydrostatic equilibrium a mean temperature,  $\theta_m$ , between these elevations, where  $\theta_m$  is defined by

$$\frac{1}{\theta_m} = \frac{1}{z_2 - z_1} \int_{z_1}^{z_2} \frac{dz}{\theta} \dots\dots\dots (11)$$

The coronal electron densities correspond to a temperature close to  $1.0 \times 10^6$  °K. Applying the equation of hydrostatic equilibrium between the levels 20,000 km. and 3000 km. at  $1.0 \times 10^6$  and  $2.71 \times 10^4$  °K. respectively, we find the mean temperature  $\theta_m$  to be  $1.24 \times 10^5$  °K. If we extrapolated the coronal electron densities down to 14,000 km. and the lower chromosphere density law up to 4000 km. we would obtain a mean temperature between these levels of  $1.05 \times 10^5$  °K.

## VI. THE H $\alpha$ LINE

An important fact which needs explanation is the formation of strong absorption lines by a chromosphere whose temperature is much higher than that of the photosphere responsible for the continuous spectrum.

Consider the central intensity of H $\alpha$  at the centre of the sun's disk. The energy passing per second through unit projected area from solid angle  $d\Omega$  in a frequency range  $d\nu$  is given by (11)

$$R_\nu d\nu d\Omega = d\nu d\Omega \int_0^\infty P_\nu \tau \alpha_{\nu,x} \exp(-\int_0^x \alpha_{\nu,x} dx) \cdot dx \dots\dots\dots (12)$$

where  $P_\nu$  is the energy emitted per second by a black body at a temperature  $\tau$  and  $\alpha_{\nu,x}$  is the absorption coefficient at a distance  $x$  from the point considered. Tables of the effective temperature  $\tau$  and of the absorption coefficient  $\alpha$  at the centre of H $\alpha$  have already been given (11) for a range of electron concentrations from  $10^8$  to  $10^{13}$  per cc.; a range of electron temperatures from  $7.5 \times 10^3$  to  $1.0 \times 10^5$  °K., and in the presence of radiation from a black body at a temperature of 5000 °K. diluted by a factor  $\kappa$ , where  $\kappa=1, 0.5$ , or 0. A significant result was an initial decrease in  $\tau$ , the effective temperature, with increase in electron temperature. This happens because electron collisions result in numerous transitions from the ground state to the metastable 2S state; thus the gas is relatively a better absorber for, than an emitter of, H $\alpha$  radiation, and the effective temperature decreases.

In the regions in which H $\alpha$  and the other Balmer lines mainly originate, the Balmer radiation probably does not differ greatly from that of a black body at a temperature of 5000 °K., but there is little doubt that the Lyman radiation, arising from chromospheric emission, will greatly exceed that of a black body at 5000 °K. Suppose initially that the radiation densities correspond to 5000 °K. black-body radiation, dilution factor  $\kappa=0.5$ ; then the absorption coefficient for an electron temperature  $\theta$  of  $2.5 \times 10^4$  °K. and concentration  $N_e$  of the order of  $10^{11}$  per cc. is about  $4 \times 10^{-31} N_e^2$ . Thus

$$\int_z^\infty \alpha dz = 8.3 \exp(-1.2 \times 10^{-8} z).$$

Roughly, the radiation can be regarded as coming from that region where

$$\int_z^\infty \alpha dz \approx 1, \text{ i.e. from a height } z \approx 1.8 \times 10^3 \text{ km. At this height, } N_e = 1.7 \times 10^{11}$$

per cc., and the effective temperature  $\tau$  is about  $3.5 \times 10^3$  °K. More accurately, we may carry out a numerical integration over all depths to obtain the central intensity of H $\alpha$ . The result is an effective temperature  $3.58 \times 10^3$  °K.

Now suppose that near L $\alpha$  and L $\beta$  the radiant energy densities are those of a black body at  $10^4$  °K., but near the Balmer lines the radiation densities are those of a black body at  $5 \times 10^3$  °K. The corresponding absorption coefficient has been computed, as in (11), for  $N_e = 10^{11}$  per cc.,  $\theta = 2.5 \times 10^4$  °K., with the result  $\alpha = 4.6 \times 10^{-9}$  cm.<sup>-1</sup>, being little affected by the increase in ultraviolet radiation. The effective temperature for the emission of H $\alpha$  is found to be  $4.67 \times 10^3$  °K., which is rather higher than for the lower Lyman radiation densities.

For both the lower and higher Lyman radiation densities assumed above, the emitted H $\alpha$  radiation corresponds to a lower temperature than that of the incident H $\alpha$  radiation. Moreover, inspection of Table 3 in (11) shows that as the density of incident H $\alpha$  radiation is reduced the effective temperature of H $\alpha$  falls, thus providing for a continuous reduction in the intensity of H $\alpha$  outwards from the photosphere. We cannot as yet predict the central intensity of H $\alpha$  from theory alone; but with H $\alpha$  radiant energy densities as assumed in the above examples (approximately as observed), the central intensities of H $\alpha$  are about 0.08 and 0.32 respectively of the surrounding continuum, whose brightness temperature at the centre of the sun's disk is about 6150 °K. These may be compared with Allen's observed central intensity (16) of 0.12.

At the limb, the brightness of the chromosphere, viewed at the centre of H $\alpha$ , should appear roughly constant to a height such that  $\int \alpha dx$ , taken along the line of sight, drops to unity, and should then decrease with height. Using the previous approximations, with the temperature uniform and about  $2.7 \times 10^4$  °K., and for simplicity considering only black-body radiant energy densities,  $T = 5 \times 10^3$  °K.,  $x = 0.5$ , we find

$$\int \alpha dx = 9.1 \times 10^{-4} \sqrt{2\pi r_0} \exp(-1.2 \times 10^{-8} z),$$

where  $r_0$  is the distance of the line of sight from the centre of the sun. The integral equals unity for  $z = 5.3 \times 10^8$ , i.e. at a height of  $5.8 \times 10^3$  km. above the limb. Hence if the chromosphere were at a constant temperature of  $2.7 \times 10^4$  °K. throughout, it would appear of approximately constant intensity, at the centre of H $\alpha$ , up to this height. If, however, the temperature be higher at elevations above 3000 km.,  $\alpha$  is less and the height over which the intensity is constant will be somewhat less. Since  $\alpha$  is not greatly affected by the ultraviolet radiation densities, neither is the height of the line greatly affected by our choice of radiation densities.

The heights computed here are in general accord with impressions gained by visual inspection of the chromosphere at the limb, but quantitative observations are lacking.



## VII. DISCUSSION

All observations on the hydrogen spectrum of the lower chromosphere to which our discussion may be applied indicate a uniform gas temperature close to 27000 °K. These observations include measurements of heights of the Balmer lines, the variations in intensity of these lines with elevation, and the profiles of chromospheric emission lines. Not only is the chromosphere effectively in hydrostatic equilibrium at this temperature, but also it appears to be in local thermal equilibrium, the electrons and atoms having the same mean kinetic energies.

Our discussion of observations at heights of the order of 500 km. upwards has shown that throughout this region hydrogen atoms are almost completely ionized, so that when the known predominance of hydrogen is taken into account, the concentration of hydrogen atoms is effectively identical with that of the electrons. Satisfactory deduction from observation of the chromospheric concentrations of the elements other than hydrogen must await theoretical discussion of the populations of the excited states of these elements under conditions which depart markedly from those of thermodynamic equilibrium.

It is not yet possible to deduce from observation the variation in temperature with height in the range 0–500 km., or 3000–20,000 km. Careful photometry of the H $\alpha$  line at eclipses to its maximum height (preferably measurements of the intensity at the centre of the line or even line profiles) would be of importance in connection with the upper of these two regions, for the behaviour of this line in the absence of thermodynamic equilibrium is now partially understood. Radio observations may also provide valuable information about the temperature of the upper chromosphere.

## VIII. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Physics, C.S.I.R.

## IX. REFERENCES

- (1) CILLIÉ, G. G., and MENZEL, D. H.—Harv. Obs. Circ. No. 410 (1935).
- (2) INGLIS, D. R., and TELLER, E.—*Astrophys. J.* **90** : 439 (1939).
- (3) PANNEKOEK, A.—*Mon. Not. Roy. Astr. Soc.* **98** : 694 (1938).
- (4) WILDT, R.—*Astrophys. J.* **105** : 36 (1947).
- (5) MITCHELL, S. A.—*Ibid.* **105** : 1 (1947).
- (6) MENZEL, D. H.—*Publ. Lick Obs.* **17** : 1 (1931).
- (7) MENZEL, D. H., and CILLIÉ, G. G.—*Astrophys. J.* **85** : 88 (1937).
- (8) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1** : 275 (1948).
- (9) MITCHELL, S. A., and WILLIAMS, T. R.—*Astrophys. J.* **77** : 1 (1933).
- (10) MITCHELL, S. A.—*Ibid.* **71** : 1 (1930).
- (11) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1** : 305 (1948).
- (12) ALLEN, C. W.—*Mon. Not. Roy. Astr. Soc.* **107** : 426 (1947).
- (13) BAUMBACH, S.—*Astr. Nachr.* **263** : 121 (1937).
- (14) REDMAN, R. O.—*Mon. Not. Roy. Astr. Soc.* **102** : 140 (1942).
- (15) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1** : 289 (1948).
- (16) ALLEN, C. W.—*Mem. Commonw. Solar Obs. Aust.* **1** (5) : 65 (1934).

# THE ELECTRONIC CHARGE AND THE OIL DROP METHOD

By V. D. HOPPER\*

(Plates 1-4)

[*Manuscript received October 4, 1948*]

## Summary

A study has been made of Millikan's oil drop method for determining the electronic charge. Owing to the uncertainty in the value of the viscosity of air, the value of the ratio  $e^{2/3}/\eta_{23}$  has been determined and quoted rather than the electronic charge. Some of the factors that have been found to influence this determination are (1) the purity of the air; (2) the influence on the field of the hole in the top plate and the layer of oil which collects in the lower plate of the condenser; (3) corrections to Stokes's equation due to the influence of the walls of the apparatus; (4) oxidation of the oil; and (5) impurities in the oil.

From a study of 47 drops of butyl sebacate, and using Stokes's analysis for the resistance to a sphere moving in a viscous medium, the value

$$e^{2/3}/\eta_{23} = (335.75 \pm 0.11) \times 10^{-5} \text{ (} e \text{ e.s.u., } \eta_{23} \text{ c.g.s.)}$$

has been obtained. This will be increased slightly if Oseen's modification to Stokes's analysis is applied, but the increase will be less than +0.04 per cent.

## I. INTRODUCTION

Millikan in 1913(1) and 1917(2) published values for the electronic charge following many years of brilliant experimental work using his well-known oil drop method. In 1930(3) he revised his 1917 value to allow for corrections in the value of the velocity of light and of the absolute ohm, giving for the electronic charge the value

$$e = (4.770 \pm 0.005) \times 10^{-10} \text{ e.s.u.}$$

In 1928 Bäcklin(4), using the X-ray method, derived a value of  $e(4.793 \times 10^{-10} \text{ e.s.u.})$  from the ratio  $F/N$  where  $F$  is the Faraday and  $N$  is Avogadro's number. The higher value of  $e$  obtained in this work was confirmed by subsequent similar experiments and Birge(5) in a recent analysis of data from this method obtained a weighted mean for the electronic charge of

$$e = (4.8021 \pm 0.0006) \times 10^{-10} \text{ e.s.u.}$$

The discrepancy between the X-ray and the oil drop values led to the belief that the difference might be due to an error in the value of the viscosity,  $\eta_{23}$ , of air at 23° C. ( $1822.6 \times 10^{-7} \text{ c.g.s.}$ ) used by Millikan. Several redeterminations

\* Physics Department, University of Melbourne.



of this constant were made and from a table of values prepared by Birge(5) the following recent determinations have been extracted (Table 1).

TABLE 1

Method	Author	$\eta_{23} \times 10^7$	Author's Claimed Uncertainty $\times 10^7$
Rotating cylinder	Bearden (1939) .. .. .	1833.79	0.06
	Houston (1937) .. .. .	1829.13	4.5
	Kellström (1937, 1941) ..	1835.00	3.0
Capillary tube	Bond (1937) .. .. .	1834.34	0.8
	Rigden (1938) .. .. .	1829.96	0.7
	Banerjea and Plattanaik (1938)	1833.75	2.2

With the range of values of  $\eta_{23}$  given in the table, Millikan's value of  $e$  ranges from  $4.796 \times 10^{-10}$  to  $4.819 \times 10^{-10}$  e.s.u. The discrepancy between Millikan's oil drop and the X-ray value could therefore readily be due to the error in the value of the viscosity of air.

Following Millikan's classical measurements a number of workers have repeated this experiment. Owing to the uncertainty in the value of the viscosity of air these results are best compared by considering the ratio  $e^{2/3}/\eta_{23}$ , which does not depend on the absolute value of the viscosity.

Table 2 gives the ratios obtained using the published data of four groups of experimenters.

TABLE 2

Author	$e^{2/3}/\eta_{23}$ (e e.s.u., $\eta_{23}$ c.g.s.)
Millikan, 1930(3) .. .. .	$334.96 \times 10^{-5}$
Bäcklin and Flemberg, 1936(6) .. ..	$334.12 \times 10^{-5}$
Ishida, Fukushima, and Suetsugu, 1937(7)	$336.65 \times 10^{-5}$
Hopper and Laby, 1941(8) .. .. .	$335.09 \times 10^{-5}$

In view of the large variations in these values, in spite of the internal consistency of separate experiments, it seemed desirable to make a careful study of this method to find, if possible, the reasons for these differences. As the experimental work proceeded the following factors were found to be significant :

- (1) The purity of the air.
- (2) The influence on the field of the hole in the top plate and the layer of oil which collects on the lower plate of the condenser.

- (3) Corrections to Stokes's equation owing to the influence of the walls of the apparatus.
- (4) The oxidation of the oil.
- (5) Impurities in the oil.

From a study of 47 drops a value of

$$e^{2/3}/\eta_{23} = (335.75 \pm 0.11) \times 10^{-5} \text{ (e e.s.u., } \eta_{23} \text{ c.g.s.)}$$

was obtained. If the recently suggested value of  $e [(4.8024 \pm 0.0005) \times 10^{-10} \text{ e.s.u.}]$  given by Du Mond and Cohen(9) is accepted, the value of the viscosity of air derived from this ratio is:

$$\eta_{23} = (1826.51 \pm 0.6) \times 10^{-7} \text{ c.g.s.}$$

The above value of  $e^{2/3}/\eta_{23}$  has been derived assuming Stokes's(10) analysis as a basis. Had Oseen's(11) modification\* of Stokes's analysis been used the above value of  $e^{2/3}/\eta_{23}$  would be increased slightly, but at the most the increase would be +0.04 per cent.

## II. DESCRIPTION OF EXPERIMENT

### (a) Apparatus

The method used for this experiment was similar to that of Millikan in that the velocities of an oil drop were measured when acted on by the earth's gravitational field and by the combined action of the gravitational field and an almost vertical electric field. The velocities were measured by illuminating the drop at intervals of 1/5 second and photographing it with a fixed camera.

The apparatus was essentially that used for the study of the effect of walls described in a previous paper(13). Oil drops were produced in an upper compartment by moving a rod against the bristles of a steel wire brush on which a quantity of oil was placed. At the bottom of this compartment was a hole through which some drops fell into a lower chamber. The lower chamber could be shut off from the upper one by a tap controlled by a long handle projecting through the heavy copper shield covering the apparatus.

In the lower chamber was placed a parallel plate condenser having its plates horizontal and consisting of two optically polished plates of kite steel, 3.8 cm. in diameter with surfaces of flatness of the order of 1/100,000 of an inch. These were separated by glass separators which had been cut from a plate of parallel walled optically flat glass. The top plate was maintained at earth potential, as were the walls of the apparatus, and a high potential was applied periodically to the bottom plate by means of a quick acting electromagnetic switch. The hole in the top plate through which the oil drops passed was 0.368 mm. in diameter and was displaced from the centre of the plate by about 2 mm.

\* Stokes made the assumption that the inertia terms of the form  $\sigma u^{\delta u} / \delta x$  ( $\sigma$  is the density of the medium) vanish in comparison with the viscosity terms of the form  $\eta \Delta u$ . Oseen has pointed out that without this assumption Stokes's equation must be modified to the form  $W = 6\pi\eta av(1 + \frac{3}{16}R)$ , where  $R$  (Reynolds's number)  $= 2av\sigma/\eta$ , the sphere assumed to be in a medium of infinite extent. There is some doubt of the preference of this equation over that of Stokes when  $R < 1$  for spheres falling in vessels bounded by walls not too remote(12).

The drops were observed with a telescope and photographed simultaneously with a camera. Both instruments contained Zeiss lenses (Planar 1:4.5,  $f=3.5$  cm.) and were mounted on a long V block which extended on opposite sides of the lower chamber. Light from the drops, before entering these instruments, passed through windows which were made of optical glass, the surfaces of which were optically flat. A photograph of the apparatus is shown in Plate 1.

### *(b) Illumination of Drops*

The source of illumination was a spark produced between two tungsten electrodes and which was triggered by a third electrode mounted vertically between them. The triggering voltage was supplied by a multivibrator unit which was designed to step down a 1000-cycle quartz oscillator control signal (correct to 1 part in  $10^6$ ) to 5 pulses per second. For visual observation of the drops the potential across the two horizontal electrodes was 2000 volts and the capacity 16  $\mu$ f, but when the photographs were taken, the voltage was slightly increased and an additional 8  $\mu$ f added.

To avoid communicating any vibration from the spark to the apparatus the electrodes were mounted on a separate platform which rested on the floor of the room, the table supporting the apparatus being mounted on concrete pillars which were built up from the ground and isolated from the building.

The spark was produced between a condensing lens and a concave reflecting mirror and the light was focused by additional lenses on to the falling oil drops. A concave mirror was attached to the inner wall of the chamber and this reflected the light back on to the droplets, thus illuminating them almost equally in opposite directions.

### *(c) Focusing of Drops*

To ensure that measurements of the velocity of the drop were made when the drop was in sharp focus the following method was adopted. The hole in the top plate of the condenser was arranged so that the drop was slightly out of focus on entering the field of view of the camera. The condenser plates were slightly tilted in the direction of the optic axis so that under the repeated action of the fields a drop was gradually brought into and finally out of focus. The plates were also tilted in the direction at right angles to the optic axis so that the path of the drop took the form of a saw tooth across the field of view. From a study of the photograph of the drop the position of best focus could be obtained.

It was also possible to repeat the process by displacing the air slightly in the apparatus so that the drop was brought back to its original plane. This was done by having two tubes leading into the lower chamber to which a hand operated pump was connected. A movement of the plunger pushed air into the vessel through one tube and withdrew an equal amount through the other. By measuring the drop velocities after successive refocusing, it was found that the slight movement of the air quickly disappeared. Care was taken to close the tap leading into the top chamber before this focusing adjustment was carried out.



Plate 2 gives an example showing the path taken by a drop. The wavy sections correspond to the small translation of the drop made to bring it back into focus. Plate 3 shows a larger drop passing twice through the region of best focus.

#### (d) *Temperature Control*

The temperature of the apparatus was controlled by a warm air method(14). It consisted of a 150-ohm copper wire resistance thermometer placed near the apparatus and made one arm of a Wheatstone's bridge so that variations in temperature caused a galvanometer spot to pass over a photoelectric cell. This controlled the heating system consisting of a three-kilowatt heater situated several yards from the apparatus. Several fans circulated the air around the room, and they were directed so that very good air circulation occurred around the apparatus. This was covered with a silver plated copper shield with walls  $\frac{1}{4}$ -inch thick and the gap between the walls and the apparatus was filled with cotton wool. The outside of the copper shield and the rest of the apparatus (camera etc.) were also covered with a thick layer of cotton wool. Occasionally the spark was operated for a long period before a suitable drop appeared in the field so that very efficient air circulation and shielding were necessary. The gradual temperature rise in the room was kept within  $0.01^{\circ}\text{C}$ . per hour and no significant temperature fluctuation occurred during the period of capturing and photographing the drop. To absorb the small amount of heat in the light from the spark two water cells were used. The one further from the apparatus contained a 0.N.19 Chance radiant heat filter. The room received no direct sunlight, was darkened with blinds, and measurements were taken on dull days or at night. Care was taken to shield the apparatus from radiant heat and at night the room was illuminated by a single sodium lamp.

The temperature inside the apparatus was measured with a fine wire copper constantan thermocouple having the cold arm in a mechanically stirred water-ice mixture. The potentiometer was checked for thermal effects by short circuiting the thermocouple using an all copper switch.

#### (e) *Voltage Stabilizer*

To obtain a constant voltage of the order of 5000 volts for the plates, the circuit\* shown in Figure 1 was constructed and has proved sufficiently effective. This method is a constant current one, employing two pentode valves in series.

Assuming that a variation  $e$  in the applied voltage produces a change of current of magnitude  $i$  in  $R$ , then

$$e = i\{[R + R_K + R_{p_1} + R_{p_2}] + \mu_1 R_K + \mu_2 R_K + \mu_2 R_{p_1} + \mu_1 \mu_2 R_K\}$$

where  $\mu_1$  and  $\mu_2$  are the amplification factors,  $R_{p_1}$  and  $R_{p_2}$  the plate resistances of the 6J7G and AF7 tubes respectively. Assuming  $\mu_1 = \mu_2 = 10^3$ ,  $R_{p_1} = R_{p_2} = 10^6$  ohm,  $R = 4 \times 10^6$  ohm, and  $R_K = 10^5$  ohm, the resulting change in voltage across the output resistance  $R$  is approximately  $4 \times 10^{-5} e$ , the remainder being across the tubes.

\* I am indebted to Major K. W. Magee for suggesting this circuit.

To obtain the closest approach to this theoretical result several precautions were found necessary. As any leakage across the valves will decrease the effective resistance of the circuit the valves were debased and held vertical and inverted on an ebonite base, the leads being kept separate by filling the cavity in the valve with paraffin wax. The dry batteries were also insulated. Filament transformers were not used owing to the low resistance between the primary and secondary windings. Several wet cells were therefore necessary for the

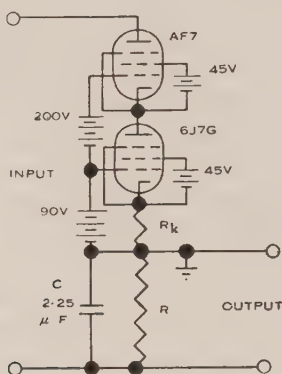


Fig. 1.—Circuit of stabilizer.

filament currents and these also were insulated. It was found necessary to prevent the occurrence of induced voltages in resistance  $R_K$  as these are amplified by the two valves. Rapid fluctuations were reduced to 1 part in 100,000 but there was usually a drop in voltage of the order of 1 part in 10,000 per hour, due mainly to filament voltage drop.

A voltage correction was necessary, varying from 1 to 5 parts in 10,000, owing to the energy needed for repeatedly charging the plates between which the oil drops fell. This charging and discharging process produced a small drop in voltage across  $C$  and the voltage reached to within 1 part in 20,000 of its final value after about fifteen seconds. For each photograph studied the average time during which the switch had been on and off was estimated, and by reoperating the switch regularly at this rate the actual applied potential on the plates during the time of photographing could be measured. To carry out this correction a simple valve circuit was set up which operated the switch automatically at the rates required. It was found that the correction was approximately proportional to the number of times the plates were charged per minute. When shielded leads were used from the condenser  $C$  to the apparatus the correction was slightly larger than for unshielded leads owing to the energy required to charge the leads.

#### (f) *Measurement of Voltage*

Resistance  $R$ , Figure 1, was made up of four 1-megohm wire wound manganin resistors and one 1000-ohm manganin resistor. The megohm resistors were



each compared with a 1-megohm substandard manganin resistor which was specially constructed to enable a direct comparison with the 1000-ohm resistor using series-parallel commutators. For the comparison the resistors were kept at 20° C. for 24 hours prior to taking measurements. To obtain the temperature coefficients of the 1-megohm resistors a current of 1.04 milliamperes was passed through them and when steady conditions had been reached the resistances were removed one at a time from the high voltage circuit and quickly compared against the substandard megohm.

The values for the 1000-ohm coil and the four megohm resistances were found to be :

	Resistance (int. ohm at 20° C.)	Temp. Coeff. (int. ohm/° C.)
Hartmann and Braun No. 2121	999,957	+0.013
Wolff 6928 .. .. .	1,000,780	+9
Pye 15200 .. .. .	1,000,395	+12
Pye 15862 .. .. .	1,000,140	+9
Pye 14919 .. .. .	1,000,468	+12
Four "megohm" coils .. .. .	4,001,783	+42

All the above resistance values are given in terms of a N.P.L. certified Cambridge 1000-ohm coil. Only the ratio of resistances enter into the derivation of  $e^{2/3}/\eta_{23}$ .

A check on the total resistance of the four megohm coils was obtained by connecting two of the resistances in parallel with the other two and measuring the resistance of the parallel combination. This gave 1,000,446 int. ohm at 20° C. That is, the total resistance in series would be 4,001,784 in agreement with the previous calibration. The resistances have been measured periodically throughout the experiment and the variation in eight months has been of the order of 1 part in 40,000. Measurements on successive days agreed to 1 part in 10<sup>6</sup>.\*

The voltage across the 1000-ohm coil was measured by means of a recently acquired Tinsley potentiometer, and this gave identical results with the Wolff Thermokraftfrei potentiometer which had previously been used.

As the voltage across the 1000-ohm coil was made very close to 1 volt it was directly compared with a standard cell† by connecting the standard cell in opposition to its voltage drop and measuring the small voltage difference. To check the voltage measurement, the standard cell was removed and the actual voltage across the 1000-ohm coil measured. This result agreed with the value obtained by the previous method.

\* I am indebted to Mr. J. L. William for his assistance in these electrical measurements and for constructing the ratio box and substandard megohm which he generously donated to the University of Melbourne.

† Three standard cells have been used and these have been calibrated for this purpose by the Defence Research Laboratories against very recently N.S.L.-certified standard cells.

*(g) Magnification of the Camera*

To obtain the magnification of the camera a similar principle to that adopted for photographing the drops was used. It consisted in placing a Grayson ruling at the focus of the camera and tilting it so that only a small section of the ruled lines were in sharp focus. From a photograph of the ruling the position of the lines which corresponded to the position of best focus could then be found and the magnification corresponding to this position measured.

The ruling (Grayson 1 cm. divided into twenty parts) was tilted about a horizontal axis perpendicular to the optic axis of the camera. This was arranged by placing it on a glass prism ( $10^\circ$  angle, base down). The visual observing telescope was removed and in its place on the V block was mounted an auto-collimating telescope. By means of this the vertical surface of the prism was rotated so that it was normal to the optic axis of the camera. The apparatus employed for rotating the prism about the vertical axis and for bringing the ruled lines into focus is shown in Plate 4. The ruling was placed so that the lines were vertical and the region of best focus normal to them. The ruling was illuminated from the side, using the same system of lenses as was used for the photography of the drops. Nine photographs of the ruling were measured, the ruling being removed and replaced several times to avoid any consistent error of setting.

The individual spacings of the Grayson ruling were carefully measured by the staff of the Metrology Section of the Defence Research Laboratories, Melbourne, and their values were used in estimating the magnification. The 1-cm. spacing of the Grayson ruling had previously been calibrated against a standard metre in this laboratory and there was exact agreement between this value and the corresponding measurement obtained by the Defence Research Laboratories.

The measurements of the magnification were corrected for emulsion shift in the following manner. The film was placed between two plates of clear glass and mounted in the plate holder. On one of these plates two sets of parallel lines 1 cm. apart were ruled, one set being approximately horizontal and the other at right angles to it. The emulsion surface of the film was pressed in contact with this ruled surface. After taking a photograph the shadow of these lines appeared on the film and by measuring the separation between the shadows of the lines the magnitude of emulsion shift could be measured. The separation between the shadows of the first and fifth horizontal or the first and fifth vertical lines of the covering graticule was measured. The actual spacings on the plate glass were 40.363 mm. and 40.275 mm. respectively, and measurements of the photographs were corrected to allow for the shift of emulsion that had occurred.

The magnification of the camera over the central three, four, and five millimetres of the field was calculated for each plate and no significant variation over these distances was evident. The mean magnification was estimated for four repetitions of each plate and the average for the nine plates gave for the magnification

$$9.3946 \pm 0.0010,$$

the error being the average deviation from the mean.

The above error does not include the error in the calibration of the Grayson ruling. Allowing for this the error increases to  $\pm 0.0025$ .

To obtain the magnification corresponding to a photograph of a falling oil drop it was necessary to measure the distance between the lines produced on the emulsion by the covering graticule. Letting  $D_H$  be the separation between the first and fifth vertical lines and  $D_V$  the separation between the first and fifth horizontal lines, the magnification  $M$  is given by

$$M = \frac{9.3946 D_H}{40.275}$$

or

$$M = \frac{9.3946 D_V}{40.363}.$$

#### (h) Variation of Field between Plates

The field  $E$  between the two plates, between which the drops fell, is calculated by the relation  $E = V/h$  where  $V$  is the potential difference applied to the plates and  $h$  is the distance between their surfaces. This expression is only true if the system approximates to two perfectly flat parallel plates of infinite extent separated by a medium of uniform dielectric constant. Factors, therefore, which may modify this expression are as follows:

- (i) The presence of the small hole in the top plate.
- (ii) The finite size of the plates (i.e. edge effect).
- (iii) The collection of oil on the surface of the lower plate.
- (iv) Flatness and parallelism of the surfaces.
- (v) Presence of local surface potentials.

These will be considered in the above order.

(i) *The Influence of the Small Hole in the Top Plate.*—A study was made of the influence of the hole on the field by studying the motion of an oil drop under the action of the gravitational and the electric field in a region covering three millimetres on each side of the vertical line passing through the hole. The method employed and results obtained have been described elsewhere(15). The results confirmed the theoretical analysis and emphasized the possibility of serious errors due to this cause.

In the apparatus used by the author the distance between the plates was 0.65695 cm. at 19° C., and to minimize the effect due to the hole, the hole was made 0.368 mm. in diameter, 2 mm. from the centre of the plate. The drop was allowed to move a millimetre or so sideways from this hole before measurements were taken. The effect on the field due to the influence of this hole was found to be negligible.

(ii) *The Edge Effect.*—The field between two semi-infinite parallel plates may be calculated and the results suggest that errors due to the edge effect for the circular plate used in the determination will be small. It was hoped that a complete solution to this problem would follow from the work of Nicholson(16) and with this in view, Dr. E. R. Love, Department of Mathematics, University of Melbourne, investigated it. Following a preliminary investigation he reported that although Nicholson's paper presented a formula for the potential due to



equal circular plates held coaxially and charged to an equal potential it seemed to involve a divergent integral. An attempt to interpret the formula led only to a contradiction. Investigation of this problem is continuing.

The field variation caused by the edge effect was also studied by Mr. J. A. Barker using the graphical method developed by Shortly *et al.*(17). An approximation was made in that the walls of the lower chamber were assumed to be symmetrically placed about the condenser plates and were at the average potential of the plates. A study was then made of the field in one-quarter of the region between the two plates using a 1-mm. net. From this work it was clear that the field variation between the plates at 8 mm. from the edge was less than 1 part in 10,000, so that the field could certainly be assumed uniform over the region studied, i.e. over a region 16 to 19 mm. from the edge.

(iii) *Collection of Oil on the Surface of the Lower Plate.*—If the surface of the bottom plate of the condenser is not cleaned, drops of oil collect on it in the region directly below the hole. This collection finally takes the form of a small raised hillock of oil. Owing to the difference in the dielectric constant of air and of oil the field vertically above the layer of oil may be appreciably increased. The lines of force would also be distorted. No measurements of oil drop velocities were made in the region below the hole where the effect would be most pronounced, and the plate was cleaned periodically.

(iv) *Flatness and Parallelism of Plates.*—The steel plates were optically polished to a flatness of 1/100,000 of an inch. The glass separating pieces were also optically polished and plane parallel. To ensure that the steel plates were in good contact with the surfaces of the glass separators, the latter were slid backwards and forwards between the steel plates until they wrung together. They were held together by means of a clamp made of polystyrene. A separating piece of the same thickness as the others was then used as a feeler gauge to check whether the plates were parallel by sliding it in at different positions near the edge. This method was checked by interferometry using a similar system in which the steel plates were replaced by semi-aluminized glass plates. Details of this are given under Section II (i).

Periodic inspections were made of the optical perfection of the surfaces of the plates and these revealed that a corrosive action had occurred on the surface of the lower plate immediately below the hole, that is, where the undeflected oil drops landed. This effect was present using Apiezon oil B as well as butyl sebacate drops. It was thus found necessary to regrind and repolish the surface of the lower plate periodically to ensure that no distortion in the field could be produced by this effect.

(v) *Surface Potential Effects.*—It has been suggested that surface potentials might possibly be influencing the fields. Difficulties had been encountered by workers on the determination of  $e/m$  using deflection methods, which were ascribed to the formation of surface charges on the plates. The effect was, however, attributed by Shaw(18) to ions formed by the electron beam and if so would not enter into this work. The possibility of surface potentials influencing the results is precluded by the fact that there is no significant difference between

the values of  $e^{2/3}/\gamma_{23}$  obtained by using potential differences of 4000 volts and 2000 volts respectively between the plates.

(i) *Separation Distance of the Condenser Plates*

The separation distance of the condenser plates was obtained by constructing a condenser using semi-aluminized plates instead of steel plates. This was set up in place of a Fabry-Perot etalon in a parallel beam of mercury light. The transmitted light was observed by a constant deviation spectrograph and the interference patterns photographed and measured. The condenser was dismantled and then reassembled and the distance remeasured. The plate glass condenser was mounted in the oil drop apparatus in the same manner as for the steel plates and after a few days removed and remeasured. A second etalon was also constructed using separating pieces cut from the same piece of parallel sided optical glass, and measurements of this have been made. The mean of all these results gave for the separation distance the value

$$0.656950 \pm 0.000006 \text{ cm. at } 19^\circ \text{ C.}$$

The glass condenser was found unsuitable for oil drop measurements owing to the corrosive action of the oil on the aluminized surface. As the steel plate condenser was constructed with similar separating pieces and in exactly the same manner there is every reason to believe that the separation distance for this condenser is within the limits given above.

(j) *Wall Correction*

The correction for the walls was mainly due to the influence of the surfaces of the condenser plates. There is a small additional correction due to the cylindrical side walls. The Lorentz formula for the single horizontal wall has been experimentally verified provided the oil drop is at a distance greater than 20 times its radius from the wall(13). Ladenburg(19) has extended the method used by Lorentz to the case of a sphere moving between two plane parallel walls and perpendicular to them. The correction factor is given as  $(1+3.3 a/h)$ , where  $h$  is the distance between the walls, when the sphere moves through the central 3/7 of the distance (this is often misquoted as the central 5/12), and as  $1+4.5 a/h$  for the central 4/5. These results have been criticized by Wessenhoff(20) on the ground that more terms should be taken. The calculation has been repeated by Miss A. M. Grant and carried out to further terms.\* It was found that this resulted in little alteration to Ladenburg's result.

In this oil drop experiment the correction factor is required over varying distances between the walls. To simplify the derivation of this correction Figure 2 was constructed. Along the axes are plotted values of  $\beta_1$  and  $\beta_2$ .  $\beta_1 h$  cm. is the distance of the drop from the lower wall when at one end, and  $\beta_2 h$  cm. its distance from the lower wall when at the other end of its measured displacement.  $(\beta_1 h - \beta_2 h)$  is the displacement of the drop over which the correction factor is required.

\* An error was found in the value given for the velocity produced by the third reflection term ( $w_3$ ) given by Ladenburg but this also made little difference to his final result.



To each curve on the diagram corresponds a value of  $k$  in the correction factor  $(1+ka/h)$ . The value of  $k$  corresponding to any pair of values  $\beta_1, \beta_2$  is obtained by selecting the curve passing through the point  $\beta_1, \beta_2$ .

The value for the correction factor  $\gamma$ , due to the cylindrical walls, is obtained by substituting in the expression  $\gamma=(1+2\cdot1a/l)$  where  $l$  (3 cm.) is the radius of the cylinder.

The application of the wall correction considerably improved the internal consistency of the results of this experiment.

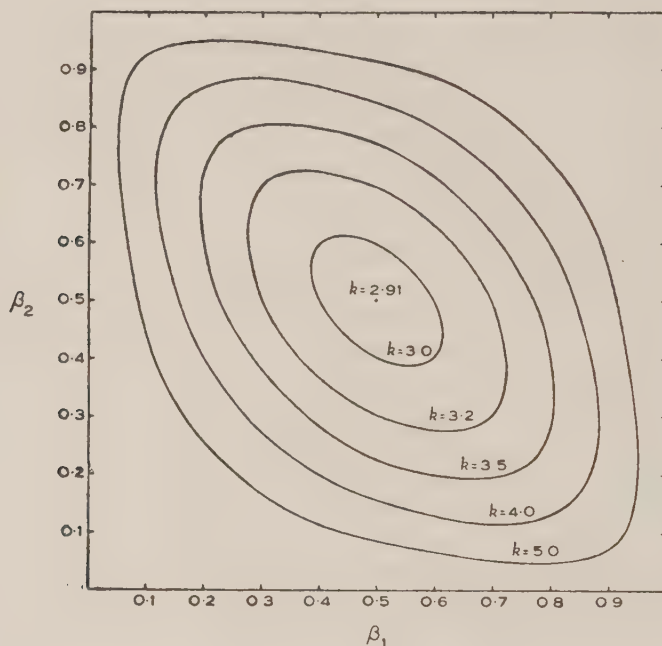


Fig. 2.—Wall correction.

#### (k) Variation of Drop Velocity with Time

Measurements on the plates which were obtained using Apiezon oil B revealed that there was a gradual variation in velocity of the drops with time when under the action of the gravitational field as well as under the action of the gravitational and electric fields. The results for drops of Apiezon oil B are shown in Figures 3A and 3B. It will be observed that over two minutes there is an increase in the distance  $S_1$  moved downwards by the drop in an interval of  $1/5$  second when under the action of the gravitational field alone, and a decrease in distance  $S_2$  under the action of the gravitational and electric fields. The effect could be explained if it was assumed that the radius of the drop increased steadily and its density changed slightly. These results are in general agreement with those obtained by Ishida, Fukushima, and Suetsugu(7) with various oils. They suggested that oxidation and absorption were responsible for the effect.

The experiment was repeated with castor oil drops and the same effect was observed but to a lesser extent. Observations for a drop of castor oil are shown in Figure 3C.

Measurements of  $S_1$  and  $S_2$  for drops of butyl sebacate,\* however, were very constant and a typical set of such measurements is shown in Figure 3D.

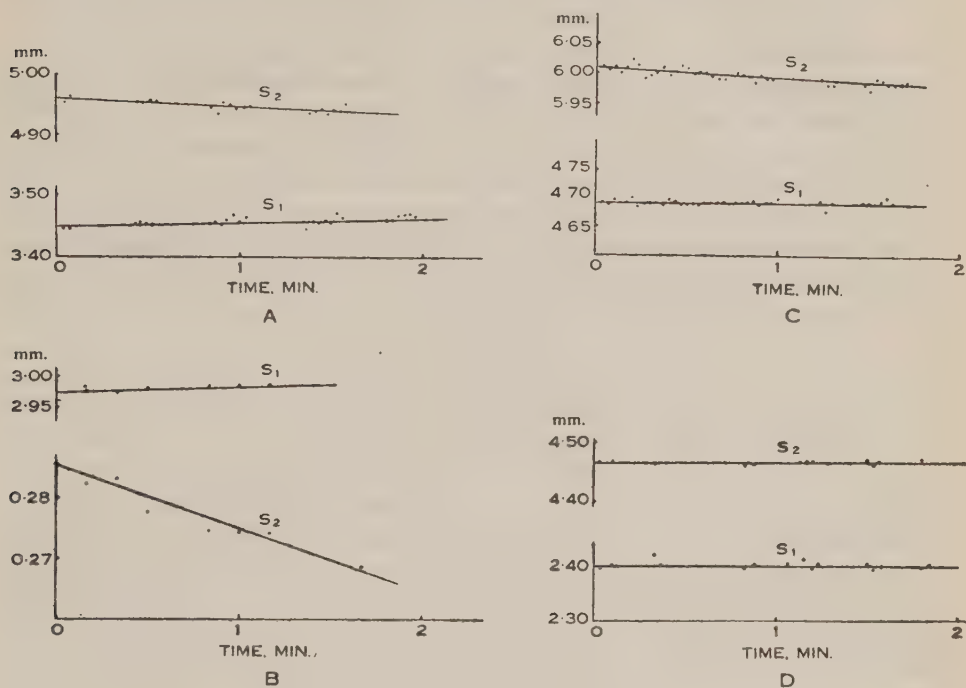


Fig. 3.—Variation of drop velocity with time.

(A) Apiezon oil B. (B) Apiezon oil B. (C) Castor oil. (D) Butyl sebacate.  $S_1$  is proportional to the velocity of the drop when under the action of gravity.  $S_2$  is proportional to the velocity of the drop when under the combined action of the gravitational and electric fields.

It seemed likely from the experiments above that the explanation of the effect observed was due to oxidation of the oil, and to ensure that the experimental technique had not changed during the course of these measurements, the brushes were carefully cleaned and a fresh supply of Apiezon oil B placed on them. Results again revealed the variation similar to that observed previously. Measurements of the drop described in Figure 3B and shown in Plate 2 were taken at this later period. Although there is a large decrease in velocity in two minutes under the combined action of gravitational and electric fields the increase in velocity under the action of the gravitational field practically compensates its effect in the evaluation of  $e^{2/3}/\eta_{23}$ . As it usually took the order of two minutes from the moment of

\* I am indebted to Associate Professor W. Davies, who, in a helpful discussion, suggested oils which might be studied.

spraying to the time of photographing the drop, it was thought that the variation in density might produce a significant though small error in the determination of  $e^{2/3}/\eta_{23}$  if this oil was used. For this reason butyl sebacate was used for the final results but the value of  $e^{2/3}/\eta_{23}$  obtained using Apiezon oil B did in fact agree with these results.

#### (l) *Purity of the Air*

It was found that if dried air, free of  $\text{CO}_2$ , was passed into the apparatus and the apparatus not re-evacuated again for several days, the value of  $e^{2/3}/\eta_{23}$  became significantly greater with the course of time (an increase of 0.3 per cent. was obtained after a few weeks). This was attributed to impurities collecting in the apparatus and contaminating the air, thus leading to a decrease in the viscosity of the gas. Day(21) in measurements of the viscosities of gases, observed a change in viscosity which was attributed to residual gases coming from the walls of the apparatus. It was therefore necessary to re-evacuate the apparatus immediately prior to taking a set of measurements.

#### (m) *Cleanliness of Oil*

The steel wire brushes were carefully washed with alcohol and when dry were thoroughly soaked with butyl sebacate. After a day's run it was found that the brush became comparatively dry and few drops were produced by stroking it. Usually the brush was not recleaned but fresh liquid was added for the further experiments. This process, however, could not be proceeded with indefinitely, as it was found that the oil on the brush became contaminated. Some of these contaminated drops have been photographed, and it was observed that the intensity of the scattered light fluctuated considerably, the appearance being quite unlike that of a normal clear droplet. Owing to this effect it was necessary to clean the brush periodically and to avoid taking observations when the brush was approaching dryness. What appeared to be fine steel particles were observed in the alcohol collected after washing the brush.

#### (n) *Density of Butyl Sebacate*

To obtain the density, a pyrex sinker was weighed in air, in water, and in butyl sebacate. For these measurements the controlling thermometer of the temperature control unit was placed near the balance case which was covered with a black cloth. Double weighings were made using calibrated weights, the temperature being read with a  $50^\circ\text{C}$ . mercury thermometer which had been calibrated against a platinum thermometer.

The following results were obtained

Temperature ( $^\circ\text{C}$ .)	Density (g.cm. $^{-3}$ )
17.01	0.93791
18.20	0.93701
21.80	0.93428

For the calculation of  $e^{2/3}/\eta_{23}$  the value of the density was taken as

$$\rho_t = 0.93701 - 0.00076_0 (t - 18.20) \text{ g. cm.}^{-3}$$

From a consideration of these measurements and others using Apiezon oil B it is concluded that the error in density is within  $\pm 0.00005$  g. cm.<sup>-3</sup>.

### III. EXPERIMENTAL PROCEDURE

During the course of the experiment the room was temperature controlled at a few degrees above the maximum daily temperature outside the building. To avoid unnecessary use of electrical energy this controlling temperature was occasionally changed, but the room temperature was allowed to reach steady conditions before measurements were made. If the control temperature was changed a delay of at least 24 hours occurred before continuing the experiment.

The procedure for a run was as follows: The voltage stabilizer set was switched on and it required about an hour's delay before reaching steady conditions. The spark voltage set was also switched on and the frequency of the multivibrator checked against the 1000-cycle supply by means of a double beam oscilloscope. Fresh oil at room temperature was dropped on to the brush and the apparatus was sealed and evacuated. The apparatus was washed out several times with dry air, free of CO<sub>2</sub>, and finally filled to the pressure required. The plate holder was next loaded with a Kodak Super Panchro Press film, which was placed between the two glass plates, the emulsion surface of the film pressing against the rulings of the covering graticule.

The temperature of the standard cells, the four megohm resistances, and the 1000-ohm resistance were recorded. The pressure of the air in the room was measured using a calibrated surveying aneroid Paulin barometer. The pressure difference between the air inside and outside the apparatus was measured with a mercury manometer. The temperature inside the apparatus was measured with the thermocouple. The voltage across the 1000-ohm resistor was measured, the potentiometer being adjusted against the corrected e.m.f. of the standard cell.

The tap connecting the lower and upper chambers was closed and drops were produced by stroking the wire brush by means of the external handle. After a few seconds the tap was opened, the spark operated at low intensity, and the drops were observed by means of the side telescope. When a drop entered the field the switch operating the potential on the lower plate was closed and the motion of the drop under the combined fields observed. If the drop moved with a convenient velocity the tap connecting the two chambers was closed and the drop moved by slightly displacing the air until it was slightly out of focus, the direction of the electric field being such as to bring it back into focus after a few excursions. The camera shutter was opened, the voltage on the spark and the capacity being increased. When the drop had finally passed through the region of focus it was again moved to the original plane and the process repeated, the drop being focused on a different region of the photographic plate. Eventually the drop reached the end of the field of view and the camera shutter was closed. It was observed visually during this period by means of the side telescope which was rotated slightly to keep the drop in view.



After the exposure was taken the thermocouple reading and the voltage across the 1000-ohm resistor were again measured. The film was developed using Kodak Special Rapid Developer at 18° C. and allowed to dry slowly in the dark.

An interval of the order of one hour elapsed between successive photographs and the above measurements were repeated in each case. Occasional measurements of the resistances were made and the rate of sparking was also checked against a 50-cycle motor operated by a tuning fork and amplifier.

The photographs were measured with a Zeiss travelling microscope and a Hilger travelling microscope. As there were often several hundred positions of a drop recorded on a photograph a complete analysis of the plate involving every position was unnecessary. To simplify matters it was decided to measure the average velocity over each rise and fall of the drop. This involved a careful measurement of two positions for each straight line of exposures. The two exposures at the ends of a row were omitted from these measurements to ensure that the drop had reached the terminal velocity. Some earlier measurements showed a slight change in velocity between the first and the second interval when the drop fell under gravity, but this was traced to the switch contact rebounding after earthing. An improvement in the design of the switch eliminated this variation.

On measuring the plates one cross-line of the eyepiece was placed at right angles to the motion of the microscope and the photographic plate was placed so that the images of the drop when falling under gravity were parallel to the microscope motion.

$S_1$ , the average distance between the successive images of the drop falling under gravity, and  $S_2$ , the component in the direction of  $S_1$  of the average distance when acted on by both fields, were measured. This was carried out over the regions of best focus and the mean values of  $S_1$  and of  $S_2$  obtained.

The position of the image of the surface of the bottom plate which was also included in the photograph, and the distance between the shadows of the first and fifth horizontal or the first and fifth vertical lines, were measured.

The inclination of the field to the vertical,  $\theta$ , was estimated from the photograph by making a measurement of the angle between the images of the drop falling vertically and the images of the drop moving under the combined action of the gravitational and electric field. The values of  $\theta$  for all photographs were less than 1°. Cosine  $\theta$  enters into the determination of  $e^{2/3}/\eta_{23}$ .

(a) Calculation of  $e^{2/3}/\eta_{23}$

The value of  $e^{2/3}/\eta_{23}$  was derived from the following equations

$$mg = \frac{4}{3}\pi a^3(\rho - \sigma)g = 6\pi\eta_1 av_1 \left(1 + k_1 \frac{a}{h}\right) \left/ \left(1 + \frac{b}{pa}\right) \right. \dots\dots\dots (1)$$

and

$$Xne \cos \theta - mg = 6\pi\eta_2 av_2 \left(1 + k_2 \frac{a}{h}\right) \left/ \left(1 + \frac{b}{pa}\right) \right. \dots\dots\dots (2)$$



where  $m$  is the mass of the drop,

$a$  is the radius of the drop,

$\rho$  is the density of the oil,

$\sigma$  is the density of the air,

$g$  is the acceleration due to gravity,

$\eta_t$  is the viscosity of the air at temperature  $t$ ,

$v_1$  is the mean velocity of the drop when under the action of the earth's gravitational field,

$v_2$  is the mean component of the velocity of the drop measured in the vertical direction when the drop is acted on by both the earth's gravitational field and the electrostatic field,

$h$  is the distance between the two plates,

$k_1$  is the coefficient of  $a/h$  to correct for the wall effect over the distance entering into the measurement of  $v_1$ ,

$k_2$  is the coefficient of  $a/h$  to correct for the wall effect over the distance entering into the measurement of  $v_2$ ,

$\gamma$  is the correction for the side walls and is equal to  $(1+2\cdot1a/l)$  where  $l$  is the radius of the chamber,

$b$  is the slip correction factor,

$p$  is the pressure of the air in the apparatus,

$X$  is the electric field,

$e$  is the electronic charge,

$n$  is the number of charges on the oil drop,

$\theta$  is the angle between the electric and gravitational fields.

Adopting the usual convention in this work of expressing the electric field in e.s.u. and all other units in c.g.s., then

$$X = \frac{V}{h} \cdot \frac{R}{r} \cdot \frac{10^8 pq}{c} \dots\dots\dots (3)$$

where  $r$  ohm is the resistance across which the voltage  $V$  international volt is measured,

$R$  ohm is the total resistance across the plates (this includes  $r$ ),

$c$  is the velocity of light ( $2\cdot99774 \times 10^{10}$  cm./sec.),

$pq$  is the conversion factor to convert international units of voltage to absolute units (1.00033).

The mean velocity  $v_1$  and the velocity component  $v_2$  are obtained from

$$v_1 = \frac{5S_1}{M} \dots\dots\dots (4)$$

and

$$v_2 = \frac{5S_2}{M} \dots\dots\dots (5)$$

where  $M$  is the magnification of the camera, and is given by

$$M = \frac{9\cdot3946}{40\cdot363} \frac{D_v}{\dots\dots\dots} \dots\dots\dots (6)$$

or 
$$M = \frac{9 \cdot 3946}{40 \cdot 275} \frac{D_H}{D_V} \dots \dots \dots (7)$$

where  $D_V$  is the distance between the shadows of the first and fifth horizontal lines and  $D_H$  is the distance between the shadows of the first and fifth vertical lines of the covering graticule.

Letting

$$S_1 \left( 1 + k_1 \frac{a}{h} \right) = S_1 \dots \dots \dots (8)$$

and

$$S_2 \left( 1 + k_2 \frac{a}{h} \right) = S_2 \dots \dots \dots (9)$$

and

$$e \left( 1 + \frac{b}{pa} \right) = e_1 \dots \dots \dots (10)$$

we have

$$a^3 = \frac{0 \cdot 75 e V \times 10^8 \times 1 \cdot 00033 R}{\pi (\rho - \sigma) g c h r} \left[ \frac{S_1 n}{S_1 + S_2} \right] \dots \dots \dots (11)$$

and

$$\frac{e_1}{\eta^{3/2}} = \frac{30 \pi h c r \eta^{3/2} (S_1 + S_2)}{n M V \times 1 \cdot 00033 R \cos \theta} \left[ \frac{S_1}{(\rho - \sigma) M} \right]^{\frac{1}{2}} \dots \dots \dots (12)$$

To determine  $S_1$  and  $S_2$  an approximate value of  $a$  was first obtained by using equation (1) and letting  $\eta_{23} = 1832 \times 10^{-7}$  c.g.s. units and  $b = 6 \cdot 0 \times 10^{-4}$  as a first approximation.

$e_1^{2/3}/\eta_{23}$  was next obtained by using Sutherland's formula for the variation of viscosity with temperature, that is,

$$\frac{\eta_{23}}{\eta_t} = \left( \frac{273 + t + C}{296 + C} \right) \left( \frac{296}{273 + t} \right)^{3/2} \dots \dots \dots (13)$$

$t$  being the temperature of the air in the apparatus, and  $C = 117$ .

#### (b) Example of Calculation

Drop No. 1 is considered as an example.

Thermocouple reading : 7·870 M.V. (From calibration data of the thermocouple against a platinum resistance thermometer the temperature of the gas was given as 19·81° C.)

Pressure : 76·177 cm.

Temperature of the standard cell which was connected to the 1000-ohm resistor ( $r$ ) : 19·90° C. (i.e. e.m.f. of standard cell 1·01825 int. volt).

Measured voltage difference between standard cell and voltage across resistor  $r$  : 0·022866 int. volt (i.e. voltage across  $r$  under steady conditions 1·01825 + 0·022866 = 1·04112 int. volt).

Drop in voltage when the switch was in operation : 0·00038 volt (i.e. voltage across  $r$  when measurements made = 1·04074 int. volt).

Temperature of 1000-ohm resistor : 19·9° C. ( $r = 999 \cdot 96$  ohm).

Temperature of four megohm resistances : (Av.) 21·3° C.

Total resistance  $R$  ("4 megohm" + "1000 ohm"): 4002838 ohm.

Density of butyl sebacate  $19.81^\circ \text{C.}$ :  $0.93579 \text{ g. cm.}^{-3}$ .

Density of air  $19.81^\circ \text{C.}$ ,  $76.177 \text{ cm. Hg}$ :  $0.00121 \text{ g. cm.}^{-3}$ .

Measurement of photograph (in mm.)

1st Set

$S_1$	$S_2$	$S_1$	$S_2$	$S_1$
43.050	44.183	43.197	37.754	37.039
15.272	15.257	12.826	13.012	11.721
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
11)27.778	7)28.926	12)30.371	6)24.742	10)25.318
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2.525*	4.132	2.531	4.124	2.532
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
$S_2$	$S_1$	$S_2$	$S_1$	$S_2$
37.055	38.488	38.955	40.118	40.094
16.421	10.718	14.220	14.825	15.333
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
5)20.634	11)27.770	6)24.735	10)25.293	6)24.761
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
4.127	2.525	4.123	2.529	4.127
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
$S_1$				
42.606				
14.684				
<hr/>				
11)27.922				
<hr/>				
2.538				
<hr/>				

2nd Set

$S_1$	$S_2$	$S_1$	$S_2$	$S_1$
42.034	42.730	44.884	43.393	45.196
14.180	13.835	11.963	14.497	14.760
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
11)27.854	7)28.895	13)32.921	7)28.896	12)30.436
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2.532	4.128	2.532	4.128	2.536
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
$S_2$	$S_1$	$S_2$	$S_1$	
46.277	46.793	46.960	45.387	
13.282	11.346	13.933	15.024	
<hr/>	<hr/>	<hr/>	<hr/>	
8)32.995	14)35.447	8)33.027	12)30.363	
<hr/>	<hr/>	<hr/>	<hr/>	
4.124	2.532	4.128	2.530	
<hr/>	<hr/>	<hr/>	<hr/>	

\* 27.778 mm. is the distance between the first and twelfth image of the drop. Values of  $S_1$  and  $S_2$  were calculated to an additional decimal place for later measurements (see Table 3).

The drop had been refocused during the exposure, the first and second set of images corresponding to the central region of focus in each case.

From these were derived

$$\text{Mean } S_1 = 2.5311 \text{ mm.}$$

$$\text{Mean } S_2 = 4.1268 \text{ mm.}$$

Corresponding to the above measurements the image of the bottom plate was at  $58.8$  mm. The average region over which  $S_1$  was measured was  $0.263h$  to  $0.737h$ , and for  $S_2$ ,  $0.274h$  to  $0.721h$  giving from Figure 2 the values  $k_1 = 3.33$ ,  $k_2 = 3.30$ .

Measurements of shadows of covering graticule  $D_H$

45.915	45.971	
5.685	5.739	
<hr/>	<hr/>	
40.230	40.232	Mean 40.231
<hr/>	<hr/>	

$$\text{i.e. magnification of camera } M = \frac{9.3946 \times 40.231}{40.275} = 9.3843.$$

The approximate value of  $a$  was found to be  $3.43 \mu$ , giving

$$S_1 = 0.25355 \text{ cm.}$$

$$S_2 = 0.41339 \text{ cm.}$$

$$\text{-----}$$

$$S_1 + S_2 = 0.66694 \text{ cm.}$$

$$\text{-----}$$

Additional data used

$$g = 979,979 \text{ cm. sec.}^{-2}$$

$$e = 4.802 \times 10^{-10} \text{ e.s.u. (assumed for calculation of } a)$$

$$l = 3 \text{ cm.}$$

$$\cos \theta = 0.99993.$$

Substitution in equation (11) gave  $a = 3.42789 \mu$

i.e.

$$1/pa = 38.296$$

and in equation (12)

$$\frac{e_1^{2/3}}{\eta_1} = 346.494 \times 10^{-5} \text{ (} e \text{ e.s.u., } \eta \text{ c.g.s.)}$$

Using Sutherland's equation to correct this result to  $23^\circ \text{C.}$  (equation (13)) we obtained

$$\frac{e_1^{2/3}}{\eta_{23}} = 343.560 \times 10^{-5} \text{ (} e \text{ e.s.u., } \eta \text{ c.g.s.)}$$

#### IV. RESULTS

Fifty-one values of  $e_1^{2/3}/\eta_{23}$  obtained from a study of forty-seven drops are given in Table 4 (see also Fig. 4). These were obtained on nine separate days from April to June 1948 and have been tabulated in the order of occurrence.

TABLE 3  
MEASUREMENTS FROM PLATES

Date	No.	Values (mm.)
April 1	1	$S_1$ 2·525, ·531, ·532, ·525, ·529, ·538, ·532, ·532, ·536, ·532, ·530
		$S_2$ 4·132, ·124, ·127, ·123, ·127, ·128, ·128, ·124, ·128
	2	$S_1$ 8·200, ·203, ·197, ·203, ·203
		$S_2$ 4·981, ·977, ·979, ·978, ·988
	3	$S_1$ 4·050, ·041, ·039, ·036, ·044, ·048, ·048, ·046, ·056, ·051, ·052, ·038, ·045, ·042, ·047
		$S_2$ 9·677, ·684, ·663, ·685, ·676, ·687, ·684, ·687, ·695, ·701, ·665, ·676, ·677, ·665
	4	$S_1$ 2·864, ·860, ·868, ·868, ·861, ·871, ·867, ·865
		$S_2$ 1·346, ·345, ·348, ·344, ·348, ·350, ·343, ·347, ·345
	5	$S_1$ 4·601, ·596, ·607, ·598, ·606, ·594, ·602, ·613, ·606
		$S_2$ 10·079, ·063, ·087, ·089, ·074, ·089, ·103, ·094, ·095
April 6	6	$S_1$ 7·217, ·207, ·202, ·205, ·213, ·209, ·220, ·213
		$S_2$ 5·878, ·892, ·873, ·874, ·870, ·871, ·879
	7	$S_1$ 5·704, ·700, ·691, ·702, ·693, ·691, ·702, ·703, ·694, ·697
		$S_2$ 10·504, ·478, ·487, ·496, ·493, ·499, ·497, ·496, ·473
	8	$S_1$ 6·795, ·797, ·781, ·789, ·787
		$S_2$ 3·004, 2·999, ·996, ·994, ·989, ·997
	9	$S_1$ 2·1696, ·1725, ·1691, ·1704, ·1727, ·1714, ·1707, ·1722
		$S_2$ 8·141, ·148, ·132, ·128, ·143, ·121, ·137, ·129, ·136
	10	$S_1$ 5·9898, 6·0010, 6·0040, 5·9927
		$S_2$ 1·8010, ·7973, ·7955, ·7984
April 9	11	$S_1$ 7·015, ·012, ·028, ·029, ·037, ·020, ·029, ·033
		$S_2$ 2·3780, ·3736, ·3772, ·3796, ·3782, ·3787, ·3799
	12	$S_1$ 6·633, ·635, ·645, ·651, ·632, ·643, ·652
		$S_2$ 6·933, ·934, ·941, ·951, ·939, ·932, ·940
	13	$S_1$ 3·2827, ·2817
		$S_2$ 6·1266, ·1408
	14	$S_1$ 2·6901, ·6873, ·6910, ·6866, ·6791, ·6904, ·6860, ·6910, ·6877, ·6900, ·6859, ·6814, ·6910
		$S_2$ 3·5845, ·5852, ·5908, ·5947, ·5831, ·5877, ·5854, ·5933, ·5860, ·6003, ·5900
	15	$S_1$ 4·7764, ·7698, ·7740, ·7790, ·7798, ·7775, ·7794, ·7843, ·7793, ·7727
		$S_2$ 4·6320, ·6427, ·6438, ·6426, ·6301, ·6370, ·6408, ·6273, ·6391, ·6282
May 13	16	$S_1$ 4·363, ·371, ·372, ·364, ·373
		$S_2$ 5·689, ·683, ·687, ·696, ·680, ·677, ·691, ·684
	17	$S_1$ 2·9086, ·9048, ·9143, ·9104, ·9038, ·9039, ·9085, ·9078
		$S_2$ 4·3858, ·3869, ·3880, ·3878, ·3935, ·3814, ·3928, ·3775
	18	$S_1$ 3·244, ·243, ·252, ·248, ·253, ·250, ·247, ·250, ·250
		$S_2$ 3·789, ·788, ·783, ·781, ·789, ·784, ·784, ·790, ·792, ·789



TABLE 3—*Continued*MEASUREMENTS FROM PLATES—*continued*

Date	No.	Values (mm.)
May 13	19	$S_1$ 2·1191, ·1202, ·1195, ·1134, ·1203, ·1167, ·1268, ·1245
		$S_2$ 13·247, ·200, ·257
		$S_2$ 13·072, ·054, ·048, ·073, ·033
	20	$S_1$ 6·284, ·284, ·279, ·277, ·283, ·279, ·273
		$S_2$ 8·069, ·086, ·078, ·076, ·082, ·077
	21	$S_1$ 1·5096, ·5039, ·5024, ·5042, ·5057, ·5057, ·5079, ·5065, ·5013, ·5065, ·5075, ·5059, ·5061
		$S_2$ 8·098, ·097, ·101, ·096, ·103, ·096, ·104, ·104
		$S_2$ 7·887, ·889, ·881, ·877
	22	$S_1$ 0·9681, ·9697, ·9686, ·9669, ·9726, ·9688
		$S_2$ 3·699, ·710, ·706, ·712, ·704, ·698
May 19	23	$S_1$ 3·3947, ·3941, ·3925, ·3910, ·3965, ·3987, ·3953
		$S_2$ 1·5938, ·5906, ·5964, ·5980, ·5958, ·6005, ·5989, ·5952, ·5971
	24	$S_1$ 2·7160, ·7173, ·7092, ·7152, ·7204, ·7184
		$S_2$ 9·102, ·104, ·098, ·097, ·104
	25	$S_1$ 4·483, ·479, ·478, ·473, ·467, ·480, ·470, ·464
		$S_2$ 9·005, 8·995, 9·007, ·004, 8·983, ·993, ·982
	26	$S_1$ 2·8077, ·8025, ·8008, ·8033, ·8000, ·8003, ·8044, ·8062, ·8011, ·8017
		$S_2$ 3·4857, ·4873, ·4795, ·4830
		$S_2$ 3·6455, ·6379, ·6511, ·6406, ·6463, ·6424
	27	$S_1$ 1·8800, ·8760, ·8823, ·8773, ·8810, ·8732, ·8805, ·8770, ·8725, ·8734, ·8718, ·8771, ·8816, ·8823, ·8761, ·8773
		$S_2$ 8·182, ·181, ·202, ·196, ·209, ·205, ·186, ·197, ·180, ·197, ·188, ·188, ·176, ·189, ·190, ·192
	28	$S_1$ 2·2993, ·2991, ·2970, ·3024, ·3047, ·2999, ·2958
		$S_2$ 5·8924, ·8935, ·8862, ·8878, ·9007, ·8795
May 20	29	$S_1$ 2·9132, ·9084, ·9141, ·9135, ·9192, ·9152, ·9274, ·9171, ·9231
		$S_2$ 9·692, ·705, ·708, ·703, ·703, ·703, ·715
	30	$S_1$ 2·4559, ·4550, ·4612, ·4633, ·4595, ·4543, ·4609, ·4576, ·4531, ·4544, ·4564, ·4568, ·4579, ·4511, ·4605, ·4572, ·4576, ·4621, ·4580, ·4610, ·4575
		$S_2$ 10·532, ·517, ·525, ·519, ·514, ·495, ·501, ·496, ·498, ·492, ·497, ·526, ·500, ·507, ·495, ·501
	31	$S_1$ 2·802, ·801, ·797, ·804, ·802, ·800, ·796, ·804, ·795, ·798, ·799, ·802, ·798, ·800, ·805, ·795, ·800, ·799, ·796, ·794, ·803, ·800, ·806, ·794, ·796, ·799
		$S_2$ 7·256, ·250, ·261, ·261, ·262, ·264, ·254, ·243, ·265, ·252, ·251, ·274, ·278, ·262, ·263, ·259, ·261, ·250, ·270, ·254, ·252, ·255, ·259, ·259, ·250

TABLE 3—Continued

MEASUREMENTS FROM PLATES—continued

Date	No.		Values (mm.)	
May 20	32	$S_1$	2·2026, ·1995, ·1992, ·2032, ·2024, ·2018, ·2050, ·2000, ·1990, ·1985	
		$S_2$	2·2545, ·2540, ·2513, ·2442, ·2468, ·2529, ·2446, ·2451, ·2542, ·2553, ·2503, ·2495, ·2498, ·2489	
	33	$S_1$	4·224, ·221, ·220, ·215, ·227, ·219, ·223, ·229, ·224, ·220, ·227, ·225, ·223, ·222	
		$S_2$	5·302, ·307, ·308, ·305, ·305, ·310, ·308, ·300, ·306, ·298	
		$S_2$	5·177, ·177, ·175, ·168	
	34	$S_1$	2·2201, ·2271, ·2285, ·2224, ·2202, ·2227, ·2233	
		$S_2$	0·4439, ·4405, ·4420, ·4392, ·4430, ·4430, ·4451, ·4436, ·4444, ·4428	
	June 9	35	$S_1$	3·3167, ·3223, ·3173, ·3148
			$S_2$	0·7037, ·7033, ·7051, ·6996, ·6996, ·7001, ·7001
		36	$S_1$	3·6433, ·6484, ·6432, ·6482, ·6528, ·6485, ·6462
$S_2$			0·6611, ·6601, ·6533, ·6520, ·6516	
37		$S_1$	2·4950, ·4981, ·4935, ·4928, ·4941, ·4935, ·4931, ·4976, ·4939, ·4866, ·4985, ·5027, ·4934	
		$S_2$	7·739, ·746, ·754, ·742, ·748, ·742, ·742, ·738, ·736, ·723, ·729	
38		$S_1$	5·2322, ·2312, ·2407, ·2325, ·2398, ·2457, ·2328	
		$S_2$	1·2022, ·2079, ·2046, ·2023, ·2020, ·2033, ·2027	
June 10		39	$S_1$	1·9838, ·9835, ·9823, ·9744, ·9718, ·9951, ·9836, ·9767, ·9793, ·9817, ·9737, ·9760, ·9878
			$S_2$	3·2420, ·2444, ·2535, ·2488, ·2439, ·2357, ·2437, ·2457, ·2460, ·2459, ·2336, ·2515, ·2456, ·2488
	40	$S_1$	3·6425, ·6511, ·6514, ·6527, ·6524, ·6549, ·6461	
		$S_2$	1·9477, ·9508, ·9494, ·9537, ·9454, ·9480	
	41	$S_1$	3·9794, ·9804, ·9937, ·9938, ·9840, ·9820, ·9837, ·9800, ·9698, ·9783, ·9743, ·9773, ·9903, ·9948, ·9862, ·9778, ·9854	
		$S_2$	2·6411, ·6464, ·6503, ·6356, ·6383, ·6408, ·6380, ·6361, ·6351, ·6408, ·6366, ·6357, ·6290, ·6456	
	42	$S_1$	3·1818, ·1717, ·1742	
		$S_2$	1·7027, ·7012, ·7023, ·7007	
	June 26	43	$S_1$	6·305, ·311, ·305, ·305, ·307, ·303, ·318, ·312
			$S_2$	1·7873, ·7858, ·7850, ·7863, ·7869, ·7864, ·7902, ·7869
44		$S_1$	1·4750, ·4652, ·4680, ·4726, ·4732, ·4712, ·4702, ·4714, ·4733, ·4705	
		$S_2$	5·7037, ·7030, ·7063, ·7037, ·7247, ·7080, ·7100, ·6953, ·7100	
45		$S_1$	1·4819, ·4830, ·4775, ·4828	
		$S_2$	3·6718, ·6757, ·6736, ·6708	
46		$S_1$	3·2920, ·2874, ·2884, ·2886, ·2906, ·2901, ·2952, ·2905, ·2886, ·2855	
		$S_2$	5·4648, ·4648, ·4656, ·4722, ·4610, ·4758, ·4715, ·4678, ·4712	
47		$S_1$	2·6355, ·6422, ·6376, ·6331, ·6420, ·6501, ·6483	
		$S_2$	0·8885, ·8888, ·8860, ·8915, ·8915, ·8929	

Only drops with charges of less than 150 electrons have been included to avoid any doubt regarding the estimation of the charge on each. Four drops changed charge at about the middle of the period of study, and two values of  $e_1^{2/3}/\eta_{23}$  have been calculated for these. One drop, on April 1, between Drop No. 4 and Drop No. 5 was omitted as it did not fit the  $e_1^{2/3}/\eta_{23}$ ,  $1/pa$  line. It seems most likely that this was due to an impurity in the oil as it was obtained near the end of the day when the brush was becoming dry. All the other drops are included in the order of appearance.

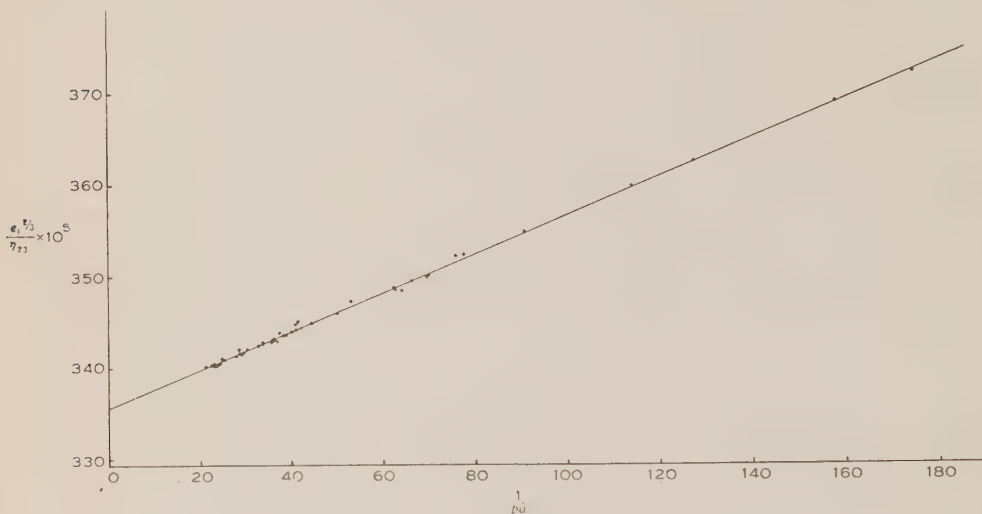


Fig. 4.— $e_1^{2/3}/\eta_{23}$  (Stokes) plotted against  $1/pa$  (butyl sebacate).

Two values of  $e_1^{2/3}/\eta_{23}$  have been tabulated for each drop. One was calculated using the method described in the example for Drop No. 1, i.e. assuming Stokes's equation as a basis. The second was obtained by multiplying this value of  $e_1^{2/3}/\eta_{23}$  by  $(1+3/16R)$  where  $R=2av\sigma/\eta$ , Reynolds's number. It seems likely that owing to the proximity of the walls the correct application of Oseen's analysis to this problem would involve a much smaller correction. The second group of values have been included to indicate the maximum possible error that could arise from this further modification of Stokes's equation.

It is assumed that the values of  $e_1^{2/3}/\eta_{23}$  and  $1/pa$  are best fitted by a straight line as predicted by the theory, i.e.

$$\frac{e_1^{2/3}}{\eta_{23}} = \frac{e^{2/3}}{\eta_{23}} + m \left( \frac{1}{pa} \right) \dots\dots\dots (14)$$

where  $m=be^{2/3}/\eta_{23}$ .

$e^{2/3}/\eta_{23}$  and  $m$  and their errors have been evaluated using least square theory for the oil drops studied at 1 atmosphere pressure and separately for the oil drops obtained at less than 1 atmosphere pressure (and one-half the field strength). Table 5 summarizes the results obtained.

TABLE 4  
DATA FOR FORTY-SEVEN DROPS OF BUTYL SEBACATE

Date	No.	$p$	$T$	$V$	$r$	$R$	$\rho - \sigma$	$M$	$\cos \theta$	$\gamma^{1/2}$	$k_1$	$k_2$	$S_1$	$S_2$	$S_1 + S_2$	$n$	$a$	$\frac{\eta_{12}}{\eta_l}$	$\frac{1}{\mu a}$	$\frac{e_1^{1/2}}{\eta_{12}} \times 10^{14}$	$\frac{e_1^{1/2}}{\eta_{12}} \times 10^{14}$
1948— April 1	176-177	19-81	1-04074	999-96	4002838	0-93458	9-3843	0-99993	1-00036	3-33	3-30	0-25311	0-41268	0-25355	0-66694	40-3	4279-1	0-00854	38-296	343-560	343-612
	276-177	19-81	1-04052	"	"	0-93458	9-3843	"	1-00065	3-34	3-37	0-82012	0-49806	0-82271	1-37236	145-6	2050-1	0-00854	21-156	340-289	340-500
	376-083	19-82	1-04034	"	"	0-93457	9-3838	"	1-00046	3-32	3-33	0-40455	0-96801	0-40541	1-37542	105-6	3435-1	0-00852	30-260	342-188	342-392
	476-083	19-83	1-04057	"	"	0-93456	9-3834	"	1-00038	3-27	3-34	0-28655	0-13462	0-28707	0-42194	27-3	6506-1	0-00849	36-004	343-076	343-076
April 6	576-083	19-86	1-04020	"	"	0-93454	9-3831	"	1-00038	3-27	3-30	0-40626	0-00859	0-46182	1-47211	120-4	6347-1	0-00840	28-359	342-067	342-218
	676-632	19-80	1-04019	999-95	4002816	0-93459	9-3797	0-99987	1-00041	3-36	3-25	0-72107	0-58705	0-72322	1-31259	135-5	8179-1	0-00860	22-430	340-413	340-593
	776-632	19-81	1-04021	"	"	0-93458	9-3797	"	1-00054	3-32	3-17	0-56977	0-40415	0-57126	1-62502	148-9	10693-1	0-00854	25-259	340-996	341-180
	876-632	19-82	1-04032	"	"	0-93457	9-3827	"	1-00053	3-33	3-24	0-67598	0-29965	0-68094	0-98142	98-5	1663-1	0-00852	23-111	340-335	340-335
April 9	976-530	19-83	1-04026	"	"	0-93456	9-3827	"	1-00053	3-31	3-33	0-21711	0-81350	0-21747	1-03228	57-3	2071-1	0-00846	24-247	345-106	345-182
	1076-530	19-83	1-04026	"	"	0-93451	9-3844	"	1-00056	3-31	3-34	0-39969	0-73871	0-70240	0-78145	73-9	2071-1	0-00831	24-664	341-195	341-324
	1175-781	21-05	1-03990	"	4002884	0-93363	9-3811	0-99992	1-00051	3-24	3-28	0-70254	0-69386	0-66601	0-94301	96-5	7494-1	0-00520	22-952	340-578	340-742
	1275-781	21-05	1-03990	"	"	0-93363	9-3856	"	1-00041	3-25	3-37	0-66416	0-69386	0-66601	0-94301	96-5	7494-1	0-00520	22-952	340-409	340-742
May 13	1375-753	21-06	1-03984	"	"	0-93363	9-3878	"	1-00041	3-23	3-41	0-38292	0-33802	0-29822	0-62876	65-3	9154-1	0-00520	33-715	342-948	342-948
	1475-753	21-06	1-03999	"	"	0-93363	9-3892	"	1-00037	3-24	3-30	0-26755	0-35802	0-29822	0-62876	65-3	9154-1	0-00509	27-883	341-409	341-512
	1575-725	21-06	1-03992	"	"	0-93360	9-3885	"	1-00050	3-37	3-35	0-47783	0-46364	0-47899	0-94375	79-4	7377-1	0-00819	29-097	341-674	341-674
	1676-073	19-96	1-04213	999-95	4002838	0-93469	9-3897	0-99990	1-00047	3-20	3-27	0-43686	0-50869	0-43735	0-73065	80-3	6766-1	0-00813	35-766	343-159	343-159
May 13	1776-045	19-96	1-04208	"	"	0-93446	9-3916	"	1-00041	3-36	3-20	0-32455	0-37869	0-32550	0-70491	48-3	8882-1	0-00815	33-858	342-768	342-830
	1875-971	19-93	1-04194	"	"	0-93447	9-3904	"	1-00041	3-36	3-20	0-32455	0-37869	0-32550	0-70491	48-3	8882-1	0-00815	33-858	342-768	342-830
	1975-920	19-96	1-04126	"	"	0-93447	9-3886	"	1-00033	3-32	3-28	0-21201	0-32347	0-21234	0-51937	84-3	1327-1	0-00813	42-050	344-522	344-646
	2075-895	19-96	1-04066	"	"	0-93447	9-3858	"	1-00037	3-34	3-41	0-62799	0-80780	0-62988	1-43937	138-5	4277-1	0-00813	24-276	340-603	340-777
May 19	2175-895	19-96	1-04072	"	"	0-93447	9-3900	"	1-00028	3-34	3-26	0-15056	0-80999	0-15076	0-96181	44-2	6338-1	0-00813	50-027	346-056	346-056
	2275-895	19-96	1-04066	"	"	0-93447	9-3839	"	1-00022	3-29	3-30	0-05691	0-37048	0-07001	0-46787	17-2	1057-1	0-00813	62-573	348-539	348-539
	2376-606	20-01	1-03695	999-95	4002825	0-93442	9-3937	0-99993	1-00042	3-51	3-57	0-33947	0-15063	0-34019	0-50016	35-3	9756-1	0-00799	32-835	342-663	342-663
	2476-530	19-96	1-03670	"	"	0-93446	9-3946	"	1-00037	3-79	3-70	0-21451	0-91014	0-27216	1-18408	74-3	5538-1	0-00813	36-768	342-914	343-014
May 20	2576-536	19-97	1-03629	"	"	0-93445	9-3909	"	1-00048	3-29	3-30	0-44732	0-80956	0-44844	1-35007	109-4	5710-1	0-00810	28-577	341-713	341-850
	2676-581	19-99	1-03624	"	"	0-93444	9-3867	"	1-00038	3-41	3-32	0-25028	0-34839	0-28080	0-62982	40-3	6093-1	0-00805	36-173	343-226	343-278
	2776-581	19-99	1-03624	"	"	0-93444	9-3881	"	1-00031	3-44	3-18	0-15775	0-81911	0-18804	0-40832	52-2	9464-1	0-00805	44-319	344-935	345-008
	2876-581	19-99	1-03624	"	"	0-93444	9-3893	"	1-00039	3-55	3-38	0-22970	0-58003	0-23038	0-82937	47-3	2653-1	0-00805	39-990	344-060	344-121
May 20	2976-327	19-96	1-03640	999-95	4002825	0-93446	9-3897	0-99990	1-00039	3-42	3-40	0-25105	0-97011	0-29921	1-24804	82-3	6839-1	0-00813	35-564	342-861	342-971
	3076-327	19-92	1-03626	"	"	0-93449	9-3897	"	1-00035	3-52	3-17	0-24574	0-79584	0-24621	0-29864	77-3	3768-1	0-00824	38-799	343-734	343-841
	3176-327	19-92	1-03619	"	"	0-93449	9-3882	"	1-00035	3-62	3-17	0-24574	0-79584	0-24621	0-29864	77-3	3768-1	0-00824	38-799	343-734	343-841
	3276-327	19-92	1-03622	"	"	0-93449	9-3818	"	1-00043	3-50	3-57	0-25011	0-92501	0-32049	0-44580	64-3	6085-1	0-00824	36-307	343-177	343-259
June 9	3376-327	19-92	1-03563	"	"	0-93449	9-3800	"	1-00047	3-57	3-49	0-42228	0-57019	0-42330	0-95504	75-4	4421-1	0-00824	41-010	344-195	344-226
	3476-327	19-92	1-03563	"	"	0-93449	9-3800	"	1-00047	3-57	3-49	0-42228	0-57019	0-42330	0-95504	75-4	4421-1	0-00824	41-010	344-195	344-226
	3576-327	19-92	1-03563	"	"	0-93449	9-3800	"	1-00047	3-57	3-49	0-42228	0-57019	0-42330	0-95504	75-4	4421-1	0-00824	41-010	344-195	344-226
	3676-327	19-92	1-03563	"	"	0-93449	9-3800	"	1-00047	3-57	3-49	0-42228	0-57019	0-42330	0-95504	75-4	4421-1	0-00824	41-010	344-195	344-226
June 10	3776-327	19-95	1-03613	"	"	0-93447	9-3872	"	1-00034	3-41	3-45	0-25225	0-34422	0-29272	0-26707	15-3	2072-1	0-00815	29-490	341-886	341-886
	3876-327	19-95	1-03613	"	"	0-93447	9-3872	"	1-00034	3-41	3-45	0-25225	0-34422	0-29272	0-26707	15-3	2072-1	0-00815	29-490	341-886	341-886
	3976-327	19-95	1-03613	"	"	0-93447	9-3872	"	1-00034	3-41	3-45	0-25225	0-34422	0-29272	0-26707	15-3	2072-1	0-00815	29-490	341-886	341-886
	4076-327	19-95	1-03613	"	"	0-93447	9-3872	"	1-00034	3-41	3-45	0-25225	0-34422	0-29272	0-26707	15-3	2072-1	0-00815	29-490	341-886	341-886
June 26	4176-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-925	349-925
	4276-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925
	4376-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925
	4476-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925
June 26	4576-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925
	4676-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925
	4776-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925
	4876-327	19-97	1-03748	"	"	0-93504	9-3890	0-99994	1-00041	3-13	3-34	0-30472	0-70160	0-33275	0-40293	54-3	8894-1	0-00797	69-564	349-911	349-925

\* Sto. kes.

† Oseen.



The Oseen correction  $(1+3/16R)$  is much smaller for the results at reduced pressures owing to the smaller value of the density of air. This accounts for the increase in  $e^{2/3}/\eta_{23}$  obtained from the 38 values from  $335.76 \times 10^{-5}$  to  $336.01 \times 10^{-5}$  ( $e$  e.s.u.,  $\eta_{23}$  c.g.s.) whereas the 13 values at less than 1 atmosphere pressure gave a value of  $e^{2/3}/\eta_{23}$  which increased from  $335.89 \times 10^{-5}$  to  $335.92 \times 10^{-5}$ . It is likely that had values of  $e^{2/3}/\eta_{23}$  been obtained for pressures of several atmospheres it would have been possible to determine experimentally the necessary correction. There is, however, insufficient variation between the results at 1 atmosphere pressure and for less than 1 atmosphere pressure to estimate it. The validity of the correction for the walls is confirmed by these results, the values of  $e^{2/3}/\eta_{23}$  and  $m$  at 1 atmosphere pressure and at less than 1 atmosphere pressure showing little correlation if the correction is not made.

TABLE 5

No. of Values	Stokes		Oseen	
	$e^{2/3}/\eta_{23} \times 10^5$	$m \times 10^5$	$e^{2/3}/\eta_{23} \times 10^5$	$m \times 10^5$
38	(1 atm. P.) $335.76 \pm 0.11$	$0.2074 \pm 0.0032$	$336.01 \pm 0.11$	$0.2033 \pm 0.0032$
13	(1 atm. P.) $335.89 \pm 0.18$	$0.2069 \pm 0.0018$	$335.92 \pm 0.18$	$0.2068 \pm 0.0018$
51	$335.75 \pm 0.05$	$0.2081 \pm 0.0008$	$335.89 \pm 0.05$	$0.2069 \pm 0.0008$

Of the remaining errors the most serious is due to the magnification. Including these the following value of  $e^{2/3}/\eta_{23}$  is obtained for the 51 values if Stokes's analysis is used as a basis

$$e^{2/3}/\eta_{23} = (335.75 \pm 0.11) \times 10^{-5} \text{ (} e \text{ e.s.u., } \eta_{23} \text{ c.g.s.)}.$$

This value would be increased by 0.04 per cent. if the maximum correction obtained by Oseen's analysis was applied to Stokes's equation.

## V. CORRECTIONS NECESSARY TO PREVIOUS DETERMINATIONS

In a letter to *Nature* in 1936, Bäcklin and Flemberg(6) reported a preliminary investigation of the electronic charge using the oil drop method. Although, to the knowledge of the author, no details have since been published, this work merits attention as it is the only result that seems to give a value of  $e^{2/3}/\eta_{23}$ ,  $[(334.12 \pm 0.62) \times 10^{-5}$  ( $e$  e.s.u.,  $\eta_{23}$  c.g.s.)], which agrees with both the X-ray value of  $e$  and Kellström's value for the viscosity of air. The present investigation, however, indicates that three important factors were omitted, namely, (1) the influence of the walls, (2) the effect due to the hole in the top condenser plate, and (3) the oxidation of the oil.

It is possible to apply with accuracy only the first of these corrections. Assuming that the measurements were made over the central region between the

plates, the corrected values of  $e_1^{2/3}/\eta_{23}$  are given in Table 6, and from these

$$e^{2/3}/\eta_{23} = (335.03 \pm 0.52) \times 10^{-5} \text{ (e e.s.u., } \eta_{23} \text{ c.g.s.)}$$

TABLE 6

RESULTS OF BÄCKLIN AND FLEMBERG CORRECTED FOR THE WALL  
EFFECT

Drop No.	$1/pa$	$e_1^{2/3}/\eta_{23} \times 10^5$ (e e.s.u., $\eta_{23}$ c.g.s.)
1	72.73	349.67
2	78.28	350.12
3	87.01	351.98
4	87.65	352.68
5	96.27	354.50
6	105.08	356.00
7	126.63	360.48
8	133.14	362.59
9	158.00	365.72

The original values of  $e_1^{2/3}/\eta_{23}$  given by Bäcklin and Flenberg have been multiplied by  $(1+k_1a/h)$  where  $k_1=3.97$ ,  $a$ =radius of drop,  $h$ =distance between the plates. They have also been corrected for  $c=2.99774 \times 10^{10}$  cm./sec. and for conversion factor to absolute volts, 1.00033.

Using least square theory

$$e^{2/3}/\eta_{23} = 335.03 \times 10^{-5} + 0.1995 \times 10^{-5} \frac{1}{pa}$$

$$(\pm 0.52 \times 10^{-5}) (\pm 0.0048 \times 10^{-5}).$$

(This value is uncorrected for effect of hole on the field.)

This value will be further increased if the influence of the hole in the top plate is considered. It has been estimated that had a hole 0.6 mm. in diameter been used, and had measurements been made below its centre, the value of  $e^{2/3}/\eta_{23}$  would have been increased to  $335.9 \times 10^{-5}$ .

It is evident from a study of the results that an insufficient number of drops was measured. For instance, if Drop No. 8 is omitted from the analysis the value of  $e^{2/3}/\eta_{23}$  corrected for the wall effect alone becomes  $335.47 \times 10^{-5}$  with an error  $\pm 0.38 \times 10^{-5}$  which is smaller than the error  $\pm 0.52 \times 10^{-5}$  obtained using the nine drops. This value of  $e^{2/3}/\eta_{23}$  corrected for the effect of a hole 0.6 mm. in diameter becomes  $336.4 \times 10^{-5}$ . As no details have been given regarding the size of the hole, it is not possible to correct their value with any accuracy. The use of Apiezon oil B in this work is also open to question, particularly if the drop is studied for long intervals after spraying.

The value of  $e^{2/3}/\eta_{23}$  derived by Ishida, Fukushima, and Suetsugu(7) in 1937 was

$$(336.65 \pm 0.21 \text{ (Slope } 0.1967)) \times 10^{-5}.$$

If this is corrected for the effect of the walls it becomes

$$336.95 \pm 0.21 \times 10^{-5}.$$

It will be remembered that a carefully prepared animal oil was used which showed some evidence of oxidation or absorption. The results were extrapolated to the moment of spraying to eliminate this effect. It was found that many drops showed a behaviour which was interpreted as due to change in shape of the drop during observation, and for this reason a considerable proportion (approximately one-half) of the measurements were omitted. The results of the 31 drops which were published show extremely high consistency, the maximum departure of a point of the  $e_1^{2/3}$ ,  $1/pa$  diagram differing from the estimated line by only 1.1 parts in 1000. It is, however, clear from a short report made later by Ishida(22) that a correction must be made to these data owing to an error in voltage measurements. After discussing several factors that had been found important it was stated: "In spite of all these precautions, our reported value is likely to be in a slight error on account of the voltage calibrations, unfortunately. The method of calibration was not satisfactory due to the relatively large amount of inner resistance of our batteries." The magnitude of this correction was not stated and no further reference to the work has been seen by the author. A volt box was used to measure the potentials of the batteries in small units and these were connected in series to apply to the plates of the condenser. It therefore seems likely that the actual potential used was higher than the assumed value. If this is so then the value of  $e^{2/3}/\eta_{23}$  would be decreased somewhat. An attempt is being made to obtain further information from the authors regarding their correction.\*

In the determination made by Hopper and Laby(8) the oil drops were larger than those studied by previous workers and the drops were photographed whilst moving. Vertical plates were also used in place of horizontal ones. The published results gave for  $e^{2/3}/\eta_{23}$  the value  $(335.09 \pm 0.06) \times 10^{-5}$ . This result was criticized by Birge(5), who derived the value  $(335.18 \pm 0.13) \times 10^{-5}$ . A more careful examination of the data was made following this criticism and the results were recalculated. It was found that several errors of entry\* had been made in the published tables. As stated on page 261 of the paper(8), more than one method could be used for the evaluation of the results. The method given in the example was used for the Apiezon oil drops and a combination of two methods was used for the castor oil drops. If, as was done by Birge, the same method is used for all the results, the value of  $e^{2/3}/\eta_{23} = (335.13 \pm 0.13) \times 10^{-5}$  is obtained. A higher value,  $e^{2/3}/\eta_{23} = (335.40 \pm 0.13) \times 10^{-5}$  results, however, if the castor oil drops are omitted.

\* In Table 7<sup>(8)</sup> the most serious errors of entry of data that had been used in the calculation of the results are: Drop No. 1, 14.97°C. to be entered in place of 15.27°C., Drop No. 4, 0.58860 m.a. in place of 0.58660 m.a., Drop No. 2, density 0.88962 in place of 0.89161.

One reason for the different values obtained from a single photograph using different methods of calculation is that the walls influence the motion of the drop producing a slight variation in velocity and curvature of path of the deflected drops. Faxén(23) has derived formulæ for the resistance to a sphere falling between two vertical walls (a) along the central line and (b) along a line three times as far from one wall as from the other. The formula for a sphere moving at various angles to the walls has not been derived. Only a rough estimate therefore can be made to allow for the effect of the walls and it appears that the values of  $e^{2/3}/\eta_{23}$  will be increased by at least 3 parts in a thousand. The uncertainty in the wall correction for this design of apparatus was one of the reasons for reverting to Millikan's vertical field method.

The work of Millikan is free from many of the above criticisms. The wall correction would increase the value of  $e^{2/3}/\eta_{23}$  from  $334.96 \times 10^{-5}$  to  $335.13 \times 10^{-5}$ . Mention should be made, however, of the fact that no correction was applied for a gradual variation in density of the oil after spraying and there is internal evidence that such a correction is necessary.

A summary of the above conclusions is contained in Table 7.

TABLE 7

Author	Published Value $e^{2/3}/\eta_{23} \times 10^5$	Corrected Value $e^{2/3}/\eta_{23} \times 10^5$	Remarks
Millikan(3) .. ..	334.96	335.13	No correction has been made for variation in density.
Bäcklin and Flemberg(6)	334.12	>335.0	No correction has been made for variation in density. Correction due to hole in top plate of the condenser is likely to increase this result considerably.
Ishida, Fukushima, and Suetsugu(7)	336.65	336.95	Voltage correction has not been made and this might reduce the value of $e^{2/3}/\eta_{23}$ .*
Hopper and Laby(8) ..	335.09	>336.1	Correction for the walls is uncertain.
Present investigation	—	335.75	A small increase (+0.04%) might be necessary due to the Oseen correction.

\* Since this paper was submitted for publication a reply has been received from Professor Ishida stating that the voltage correction was determined in 1940 and this decreased their 1937 value of  $e$  by 1.1 per cent. (i.e.  $e^{2/3}/\eta_{23}$  reduced from  $336.65 \times 10^{-5}$  to  $334.16 \times 10^{-5}$ ). A new determination has been carried out by Ishida, Suetsugu, and Matsui giving, it is stated,  $e = 4.794 \times 10^{-10}$  e.s.u. for  $\eta_{23} = 1830 \times 10^{-7}$  c.g.s., i.e.  $e^{2/3}/\eta_{23} = 334.72 \times 10^{-5}$ . Using the drop velocities measured half an hour after atomization and assuming no change in density, the value  $e^{2/3}/\eta_{23} = 334.95 \times 10^{-5}$  was obtained. If the same correction for the walls is assumed as for the 1937 work, these values of  $e^{2/3}/\eta_{23}$  will be increased to  $335.02 \times 10^{-5}$  and  $335.25 \times 10^{-5}$  respectively.



## VI. CONCLUSION

A study has been made of the oil drop method used for the determination of the electronic charge. Owing to the uncertainty in the value of the viscosity of air, the ratio  $e^{2/3}/\eta_{23}$  has been derived as this involves only the temperature coefficient of the viscosity of air and not its absolute value.

Some of the factors that have been found to influence this determination are :

- (1) The purity of the air.
- (2) The influence on the field of the hole in the top plate and the layer of oil which collects on the lower plate of the condenser.
- (3) Corrections to Stokes's equation owing to the influence of the walls of the apparatus.
- (4) The oxidation of the oil.
- (5) Impurities in the oil.

When errors due to the above factors were removed and Stokes's analysis was assumed the following result was obtained

$$e_1^{2/3}/\eta_{23} \times 10^5 = (335.75 \pm 0.05) + (0.2081 \pm 0.0008) \frac{1}{pa},$$

$e_1^{2/3}/\eta_{23}$  being the value of  $e^{2/3}/\eta_{23}$  at pressure  $p$  for a drop of radius  $a$ , when the correction for the slip term is omitted. The errors given above were obtained by applying least square theory to the result and do not include errors in subsidiary measurements. When allowance is made for these the following result is obtained for  $e^{2/3}/\eta_{23}$

$$(335.75 \pm 0.11) \times 10^{-5} \text{ (e e.s.u., } \eta_{23} \text{ c.g.s.)}.$$

This value would be increased by  $+0.04$  per cent. if the maximum correction obtained by Oseen's analysis(11) were applied.

The value of the electronic charge obtained by this method can be derived from  $e^{2/3}/\eta_{23}$  when the viscosity of air is more accurately known. If, however the value  $4.8024 \times 10^{-10}$  e.s.u. is accepted for the electronic charge, the value of the viscosity of air is derived as

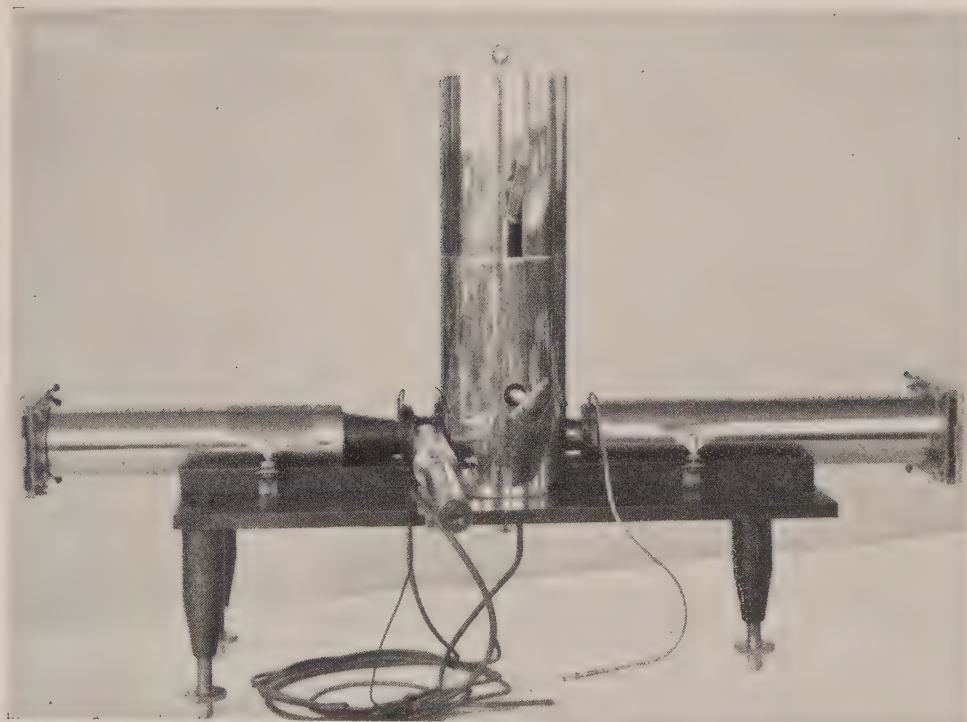
$$\eta_{23} = (1826.51 \pm 0.6) \times 10^{-7} \text{ c.g.s. units.}$$

This value is higher than the original value chosen by Millikan but much lower than several of the most recent values of this constant (see Table 1).

The other oil drop determinations, corrected in the same way and assuming the above value of the electronic charge, give derived values of  $\eta_{23}$  ranging from  $1829.9 \times 10^{-7}$  c.g.s. (Millikan) to  $1820.0 \times 10^{-7}$  c.g.s. (Ishida, Fukushima, and Suetsugu).

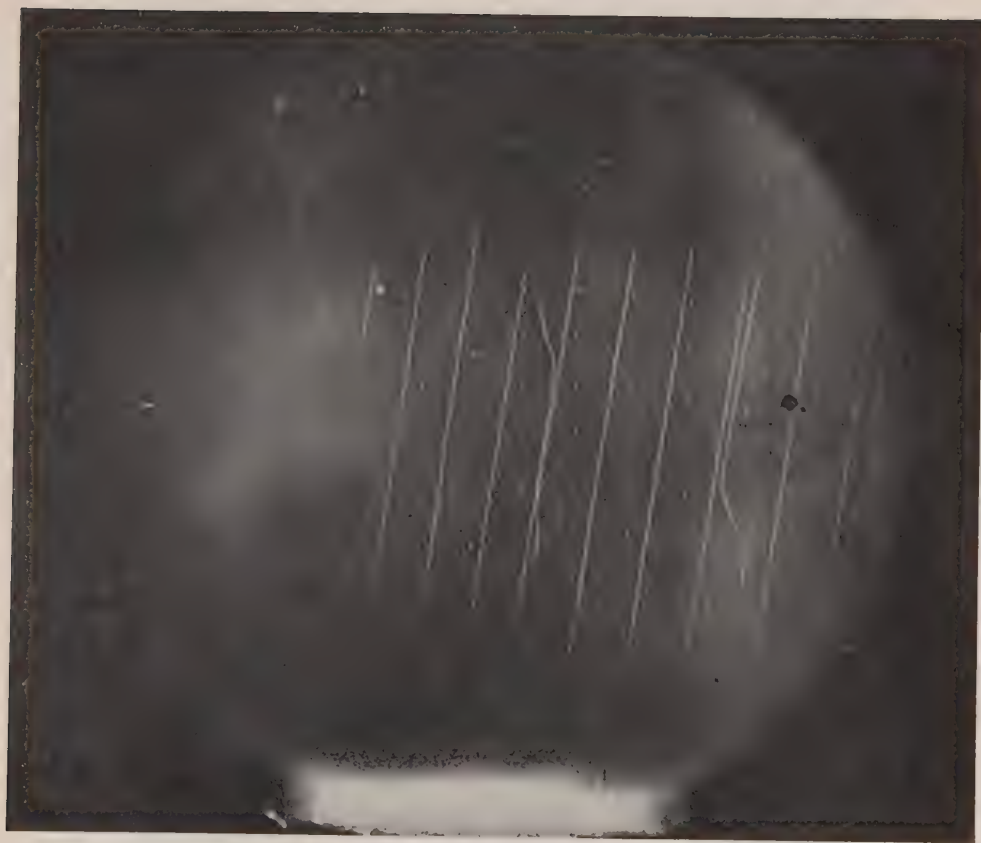
## VII. ACKNOWLEDGMENTS

I wish to express my appreciation for the valuable assistance received from many during the course of this investigation ; reference has been made to some of these in the text. In particular, the assistance of the staff of the Metrology Section of the Defence Research Laboratories and the staff of Kodak Research Laboratories (Melbourne) is gratefully acknowledged.



HOPPER.—THE ELECTRONIC CHARGE AND THE OIL DROP METHOD





HOPPER.—THE ELECTRONIC CHARGE AND THE OIL DROP METHOD

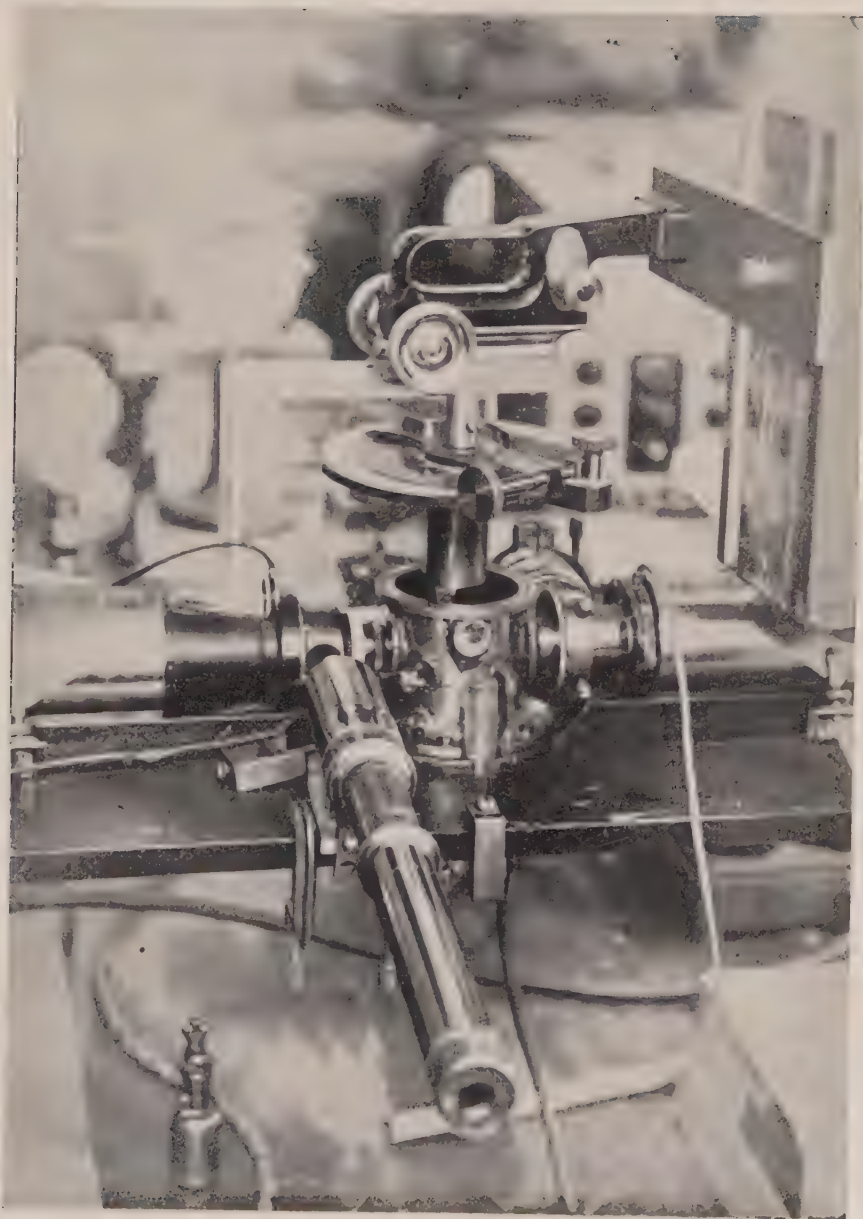






HOPPER.—THE ELECTRONIC CHARGE AND THE OIL DROP METHOD





HOPPER.—THE ELECTRONIC CHARGE AND THE OIL DROP METHOD





Miss A. M. Grant assisted in many of the measurements and in computing the results. To Professor L. H. Martin I am indebted for many helpful discussions and for continuous encouragement during the course of this investigation.

Financial assistance was provided by the Council for Scientific and Industrial Research.

### VIII. REFERENCES

- (1) MILLIKAN, R. A.—*Phys. Rev.* **2**: 109 (1913).
- (2) MILLIKAN, R. A.—*Phil. Mag.* **34**: 1 (1917).
- (3) MILLIKAN, R. A.—*Phys. Rev.* **35**: 1231 (1930).
- (4) BÄCKLIN, E.—*Inaug. Diss. Uppsala Univ. Arsskt.* (1928).
- (5) BIRGE, R. T.—*Amer. J. Phys.* **13**: 63 (1945).
- (6) BÄCKLIN, E., and FLEMBERG, H.—*Nature* **137**: 656 (1936).
- (7) ISHIDA, Y., FUKUSHIMA, I., and SUETSUGU, T.—*Sci. Pap. Inst. Phys. Chem. Res. Tokyo* **32**: 57 (1937).
- (8) HOPPER, V. D., and LABY, T. H.—*Proc. Roy. Soc. A* **178**: 243 (1941).
- (9) DuMOND, J. W. M., and COHEN, E. R.—*Rev. Mod. Phys.* **20**: 82 (1945).
- (10) STOKES, G. G.—“Mathematical and Physical Papers”, Vols. II and III (Cambridge Univ. Press, 1880).
- (11) OSEEN, C. W.—*Ark. Mat. Astr. Fys.* **6**: No. 29 (1910).
- (12) BARR, G.—“A Monograph of Viscometry”, Chap. VIII (Oxford Univ. Press, 1931).
- (13) HOPPER, V. D., and GRANT, ALISON M.—*Aust. J. Sci. Res. A* **1**: 28 (1948).
- (14) HOPPER, V. D.—*Proc. R. Phys. Soc. Lond.* **54**: 55 (1942).
- (15) HOPPER, V. D.—*Nature* **158**: 786 (1946).
- (16) NICHOLSON, J. W.—*Philos. Trans. A* **224**: 303 (1923-1924).
- (17) SHORTLEY, G., WELLER, R., DARBY, P., and GAMBLE, E. H.—*J. Appl. Phys.* **18**: 116 (1947).
- (18) SHAW, A. E.—*Phys. Rev.* **44**: 1009 (1933).
- (19) LADENBURG, R.—*Ann. Phys. Paris* **23**: 447 (1907).
- (20) WESSENHOFF, J.—*Ann. Phys. Lpz.* **62**: 1 (1920).
- (21) DAY, R. K.—*Phys. Rev.* **40**: 281 (1932).
- (22) ISHIDA, Y.—*Sci. Pap. Inst. Phys. Chem. Res. Tokyo* **33-34**: 1785 (1937-38).
- (23) FAXÉN, H.—*Ann. Phys. Lpz.* **68**: 89 (1922). *Ark. Mat. Astr. Fys.* **18**: No. 29 (1924-25).

### EXPLANATION OF PLATES 1-4

#### PLATE 1

Apparatus used for photographing drops.

#### PLATE 2

Path of Apiezon oil B drop ( $n=22$ ,  $a=3.90\mu$ ) which had been refocused several times. The wavy sections correspond to the small translation of the drop made to bring it into focus. Measurements from this course were used in preparing Fig. 3B.

#### PLATE 3

Path of more highly charged drop ( $n=96$ ,  $a=5.62\mu$ ) passing twice through the region of best focus.

#### PLATE 4

Apparatus used for mounting the Grayson ruling in the determination of the magnification of the camera.

# SOME ASPECTS OF THE GLOW DISCHARGE BETWEEN COAXIAL CYLINDERS IN THE PRESENCE OF A NON-HOMOGENEOUS AXIAL MAGNETIC FIELD

By J. M. SOMERVILLE,\* K. S. W. CHAMPION,† and E. K. BIGG\*

[Manuscript received September 13, 1948]

## Summary

The characteristics of the glow discharge in air between coaxial cylindrical electrodes in the presence of a non-homogeneous axial magnetic field are described. With the outer cylinder as cathode the discharge is narrow and takes place near the position of maximum magnetic field. If the cathode is cut transversely into halves dividing the discharge there may be a negative incremental resistance between these halves. This negative resistance may be used as the basis of an oscillator working with an efficiency of about 70 per cent. up to frequencies of about 100 kilocycles per second.

## I. INTRODUCTION

A comprehensive survey of the characteristics of the glow discharge between coaxial cylinders in the presence of a homogeneous axial magnetic field has been given by Penning‡ for the gases argon and helium. The characteristics of the similar discharge when the magnetic field is not homogeneous but is derived from a short coaxial coil are of some interest. During a study of them we found that with the arrangement of Figure 1 in which the outer cylinder is the cathode and is split transversely into two sections separated by a narrow

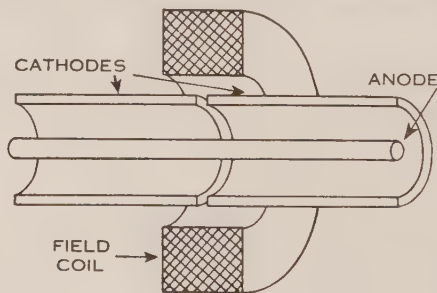


Fig. 1

gap, a strong negative incremental resistance may exist between these sections or between either section and the central anode. This negative resistance may

\* Department of Mathematics and Physics, New England University College, Armidale, N.S.W.

† Department of Physics, University of Queensland, Brisbane.

‡ Penning, F. M.—The glow discharge at low pressures between coaxial cylinders in an axial magnetic field. *Physica* 3 : 873 (1936).

be as low as three or four thousand ohms and may be used to sustain oscillations in a tuned circuit. The characteristics of an oscillator of this type are described in Section V.

## II. DESCRIPTION OF APPARATUS

The tubes were constructed with copper electrodes, the inner having a diameter of 7.7 mm. and the outer an internal diameter of 22 mm. The electrode surfaces were smooth but not specially polished and were cleaned with dilute acids, alcohol, and distilled water before assembly. The use of waxed joints precluded proper outgassing but under working conditions clean-up of gas was rapid. The tubes were about 40 cm. long and it was found that this length was sufficient to eliminate end-effects. Tubes could be assembled with cathode or anode divided transversely into two sections separated by a  $\frac{1}{2}$ -mm. gap; with one or several probes in the cathode wall; and with any combination of these features.

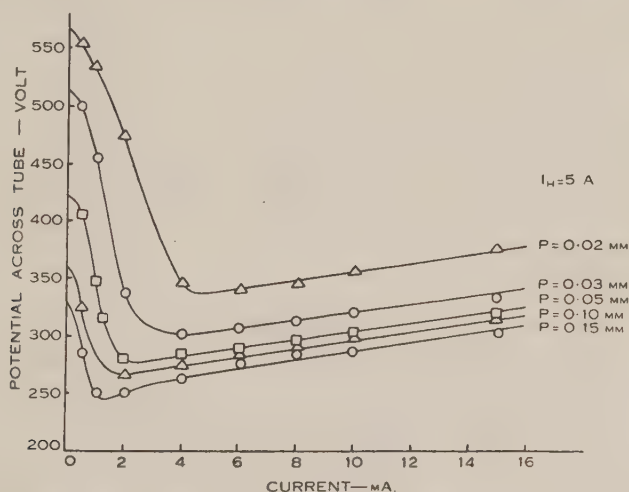


Fig. 2a.—Current-voltage characteristics of the discharge for various gas pressures.

The field coils were mounted coaxial with the tube and could be moved parallel to it. The coils consisted of 450 turns of 16-S.W.G. wire with inner diameter 5.7 cm., outer diameter 14 cm., and a width of 4.0 cm. The axial field strength at the centre of the coil was 56 oersted for 1 ampere field current ( $I_H$ ). The axial field of a single coil falls to 81 per cent. of its maximum value at a distance of 2 cm. from the central plane.

The gas used was air dried over phosphorus pentoxide. The fact that the tubes could not be outgassed meant that the constitution of the gas was subject to some variation with time as were the states of cathode and anode surfaces owing to sputtering and chemical action. Consequently the curves reproduced with this article are indicative of the general nature of the characteristics of the discharge but should not be regarded as sources of precise data. They are



consistently reproducible in general form but because of the variations in conditions mentioned above they are subject to some variation in detail.

### III. CHARACTERISTICS OF THE DISCHARGE

#### (a) *Voltage-current Characteristics*

When the tube is operated as a simple diode, with outer electrode as cathode and a single field coil, voltage-current characteristics like those of Figure 2 are obtained. For low discharge currents ( $I$ ) (less than about 1 mA.) the tube voltage falls sharply, sometimes discontinuously, with increasing current to reach a

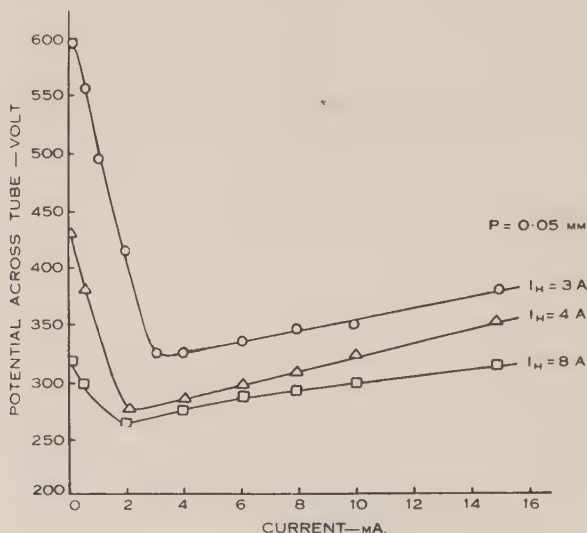


Fig. 2b.—Current-voltage characteristics of the discharge for various magnetic field strengths.

minimum for currents of a few milliamperes, thereafter rising gradually. For a given current, decreasing the pressure or the magnetic field strength increases the voltage drop across the tube.

#### (b) *Current Distribution Along the Axis*

Three ways of measuring the distribution of cathode current along the length of the tube were tried. Two methods are based on the assumption that the discharge will move bodily with the magnetic field if the latter is moved coaxially relative to the tube. They are :

(i) Using a tube with the cathode split into two sections separated by a narrow gap as in Figure 1, the magnetic field coil is moved along the length of the tube past the gap and the currents to the two sections of the cathode measured. Graphical differentiation of the curve of current to one cathode  $v$ . field position gives the linear current density at the cathode.

(ii) Using a tube with a small probe or sound at the cathode and moving the field as before, curves of probe current  $v$ . field position may be obtained in which

the probe current might be expected to be proportional to the linear current density at the cathode.

As a check, a third method was tried in which the magnetic field is not moved.

(iii) Using a tube with a number of similar cathode probes suitably placed along the length of the tube the values of probe current at different positions relative to the magnetic field may be obtained and compared with values of probe current *v.* field position obtained by method (ii). The field coil and discharge remain stationary.

*Comparison of Different Probes.*—Two types of probe were tried for use in methods (ii) and (iii) and may be called “sunk” and “flush” probes respectively. A sunk probe consists of a copper collecting electrode mounted nearly flush with the *outer* wall of the cathode and communicating with the discharge through a circular hole 3 mm. long and 1.5 mm. in diameter bored through the cathode. A flush probe consists of a copper collecting electrode

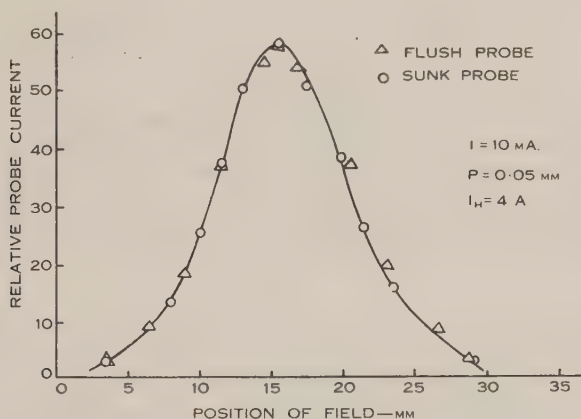


Fig. 3a.—A comparison of the currents to sunk, and flush, cathode probes for different positions of the magnetic field coil.

1.5 mm. in diameter mounted flush with the inner wall of the cathode and separated from it by an insulating gap. Probe currents were of the order of  $10 \mu\text{A}$ . As may be seen from Figure 3 (a), the agreement between current distribution curves obtained by using the two types of probe is close. The ratio of the currents to the two types of probe in the same position in the discharge is not constant but varies somewhat with pressure, magnetic field, and total current. This is not surprising since the magnetic field prevents electrons from traversing the narrow channel through the cathode to reach the sunk probe. Thus the flush probe would collect electrons emitted from the cathode and returned to it by the field while the sunk probe would not. Again, some of the positive ions which would be collected by the flush probe would be lost by diffusion before reaching the sunk probe. However, the close agreement between the shapes of the current distribution curves given by the two probes makes it

immaterial which type is used for the purpose. Because of their greater ease of construction, sunk probes were used for both methods (ii) and (iii) above.

*Comparison of Moving Field plus Single Probe with Stationary Field plus Several Probes (methods (ii) and (iii)).*—Agreement is satisfactory, as shown by the example of Figure 3 (b). There appears to be no reason to doubt that the discharge moves bodily with the field.

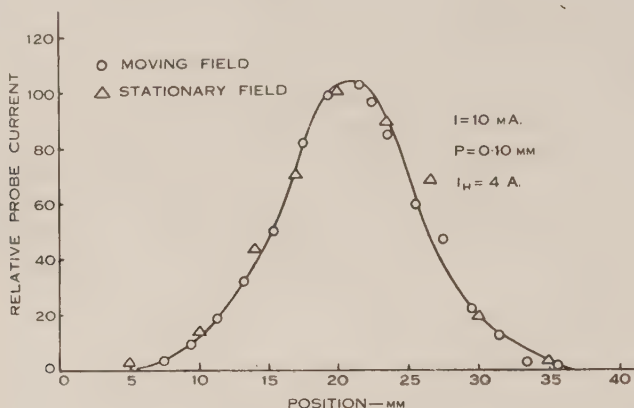


Fig. 3b.—A comparison of moving field and stationary field methods of measuring cathode current density. Circles represent values of current to a single cathode probe for various positions of a moving magnetic coil. Triangles represent values of the currents to a number of cathode probes located at various distances from a fixed field coil.

*Comparison of Probe Methods with Split-Cathode Method.*—Agreement between current density curves obtained using probes and split-cathode methods is fairly good, although accurate graphical differentiation of the cathode curves is difficult and generally speaking split-cathode current density curves are not as reproducible as those obtained with probes. An example of the agreement between probe and split-cathode current density curves is shown in Figure 4.

*Cathode Current Density Distribution.*—As a result of the comparisons made above, the method (ii) of moving the field coil past a fixed probe was chosen as the most convenient and consistent method of measuring the linear current density at the cathode, and the results of measurements made in this way for various conditions of gas pressure, total tube current, and magnetic field strength are shown in Figures 5, 6, and 7.

In connection with Figure 5 it should be noted that in the absence of a magnetic field the discharge extinguishes at a pressure of about 0.2 mm. so that a magnetic field will locate a discharge and control its width even when it is not necessary for its existence.

*Comparison of Cathode and Anode Current Densities.*—Using a tube with a split anode and a method similar to method (i) above, anode current-density measurements were made and Figure 4 shows a comparison of the cathode and

anode current density distribution. It will be seen that the linear current density at the anode is considerably greater than at the cathode and that the discharge is narrower than at the cathode. This narrowing of the discharge towards the anode may be explained as follows :

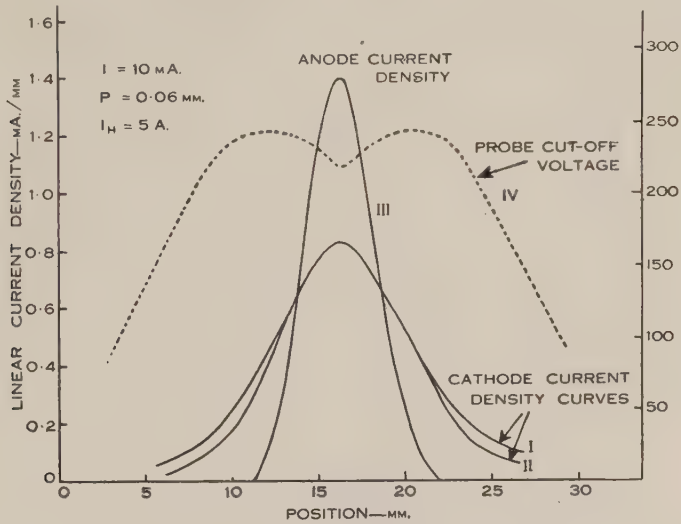


Fig. 4.—Curves I and II are those of cathode current density measured by the probe method (Curve I) and the split-cathode method (Curve II). Curve III shows anode current density measured by the split-anode method. Curve IV shows the cathode-to-probe bias voltage necessary to reduce the probe current to zero.

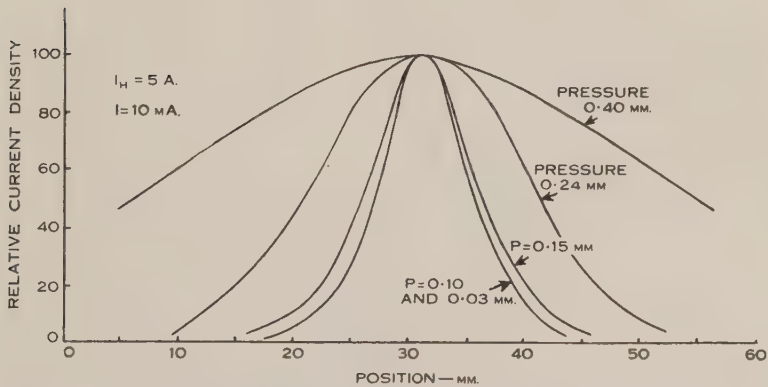


Fig. 5.—Cathode current density curves for various gas pressures. The curves for pressures from 0.1 mm. to 0.83 mm. are indistinguishable.

The lines of magnetic force have the general shape shown in Figure 8. Electrons leaving the cathode at the edges of the discharge will have on the average a component of velocity along the lines of force of the magnetic field towards the centre of the discharge and in their motion from cathode to anode will drift in that direction, producing a discharge narrower at the anode than at



the cathode. The dotted curve of Figure 4 shows the retarding potential necessary to reduce to zero the current to a sunk cathode probe in different parts of the discharge. This bias should approximate to the maximum energy of positive ions reaching the cathode in the neighbourhood of the probe. Near the centre of the discharge, the retarding potential is nearly equal to the voltage drop across the tube, showing that some ions reaching the cathode have been formed

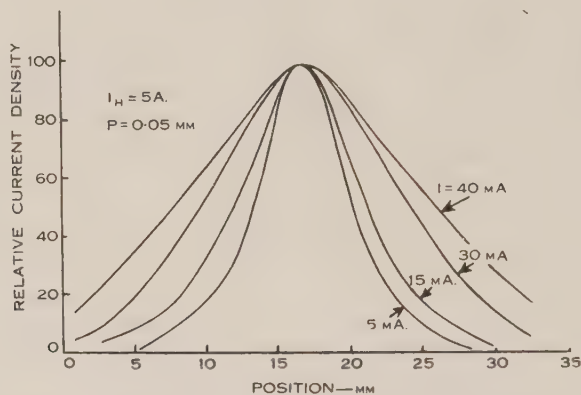


Fig. 6.—Cathode current density curves for various total currents.

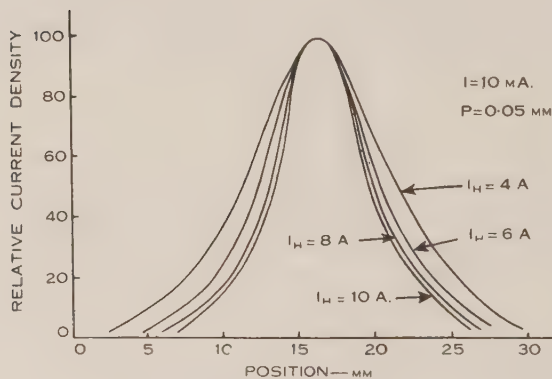


Fig. 7.—Cathode current density curves for various magnetic field coil currents.

near the anode. However, on approaching the edges of the discharge, the necessary retarding potential falls rapidly, showing that the ions of maximum energy are formed progressively nearer the cathode, consistently with the shape of the discharge shown in Figure 8.

*Current Density Curves with More Uniform Magnetic Fields.*—If the magnetic field is made more uniform for a greater length along the axis by using two field coils the discharge is only slightly broadened and as soon as the field has two maxima, the discharge locates itself near one or other of these maxima and is easily displaced from one position to the other by a small change in the current

to one of the field coils. (On one occasion a double-peaked current density distribution was observed which was very unstable. The conditions for such a distribution appear to be very critical and subsequent attempts to obtain a similar distribution failed.)

### (c) Other Features of the Discharge

The influence of the curved magnetic field in deflecting electrons to the centre of the discharge appears to play an important role in determining the structure of the discharge with internal anode. The following experiments throw light on this matter.

#### (i) Experiments with Baffles

A split-cathode split-anode tube was fitted with a thin mica baffle extending from anode to cathode and lying between the two sections of the anode and the two sections of the cathode. (Gas was able to pass round the edge of the baffle.) The discharge is then situated entirely on one side of the baffle. When the field coil is symmetrically over the baffle the discharge is unstable and switches backwards and forwards from one side of the baffle to the other.



Fig. 8.—Sketch of lines of magnetic intensity and the general shape of the discharge. (Not to scale.)

When a circular mica baffle is inserted extending from the anode to within 2 mm. of the cathode, the major part of the discharge occurs on one side of the baffle. When the magnetic field is moved along the tube the discharge does not in general follow it continuously but lags behind, and then, when the coil is well past the central position, jumps discontinuously to the other side, the process being repeated in reverse when the coil is moved across the baffle in the opposite direction. Under these conditions a symmetrical state with half the discharge on either side of the baffle cannot be achieved. If, however, the magnetic field is increased beyond a critical value, the discharge moves continuously with the field, and a symmetrical condition can readily be obtained. The critical field is of the order of magnitude required to allow electrons from the cathode to move across the plane of the baffle between the cathode and the anode.

When an annular mica baffle was inserted extending from the cathode to a distance of about 2 mm. from it, leaving a 5-mm. gap between the baffle and the anode, behaviour similar to that described in the preceding paragraph was observed, with the difference that the discharge could be made to move continuously with the field by making the magnetic field sufficiently *small*, of an order of magnitude sufficient to allow electrons from the cathode to cross the plane of the baffle in the space between the baffle and the anode.

These observations suggest that the convergence of electrons towards the centre of the discharge and the transfer of electrons from either side across the central plane are essential features of the discharge, and the discharge will follow the field past a baffle only if electrons from one side are able to reach the other to "lead" the discharge.

(ii) *Discharge with Inner Electrode as Cathode*

Reference to Figure 8 shows that if the central electrode were the cathode, electrons from the cathode would diverge from the centre of the discharge, and there would be no transfer of electrons across the central plane. We might expect the discharge under these conditions to have characteristics very different from those of the discharge with central anode. This proves to be the case and we found that with the central electrode as cathode:

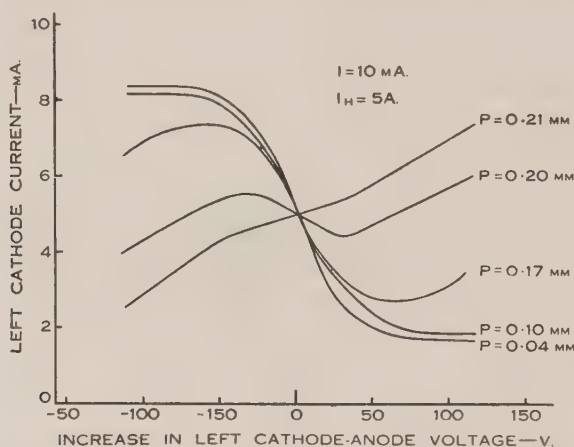


Fig. 9.—Current of left cathode section of a split-cathode tube as a function of the voltage between left cathode and anode, total anode current being constant. (Under these conditions the mean of left and right cathode-to-anode voltages is also constant.) A downward slope from left to right implies a negative incremental resistance.

(i) The discharge at the central electrode is much wider, having a total width of the order of 10 cm. as compared with about 3 cm. when the central electrode is the anode.

(ii) The discharge is wider at the anode than at the cathode.

(iii) A baffle extending completely from the anode to the cathode does not affect the discharge which moves continuously past it as the field is moved.

#### IV. NEGATIVE RESISTANCE CHARACTERISTICS OF A SPLIT-CATHODE TUBE

A split-cathode tube like that of Figure 1 exhibits negative resistance characteristics as follows: If the potential drop between the anode and the left cathode is  $V_L$  and that between the anode and the right cathode  $V_R$  and if  $I_L$  and  $I_R$  denote the corresponding cathode currents, it is found that if  $V_L$

is increased and  $V_R$  decreased keeping  $V_L + V_R$  constant, then  $I_L$  is decreased and  $I_R$  increased,  $I_L + I_R$  remaining constant. Thus the current decreases in that part of the tube where the anode-cathode potential drop is increased and vice versa.

Figure 9 shows the current to one section of the cathode as a function of the increase of potential difference between one cathode section and the anode for various gas pressures. On this graph a slope downwards from left to right corresponds to a negative incremental resistance. As the gas pressure decreases below the critical value at which the discharge will not pass without the magnetic field (about 0.20 mm.), the incremental resistance  $\Delta V_L / \Delta I_L$  changes from positive to negative.

Curves showing the change in current distribution with potential difference between the two cathode sections when the incremental resistance is negative are shown in Figure 10. These curves were obtained using a split-cathode tube with a number of probes near the cathode gap.

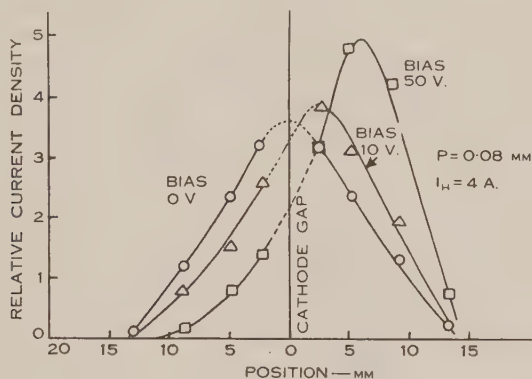


Fig. 10.—Cathode current density curves in a split-cathode tube for various biases between cathode sections, total current being constant. The magnetic field is placed centrally over the cathode gap. The larger current flows to the cathode section whose potential is nearer to the anode potential.

Figure 11 shows the values of the initial negative incremental resistance obtained for various total currents. The negative resistance does not depend much on pressure or magnetic field as long as the pressure is low enough to prevent a discharge without the magnetic field. As the pressure increases beyond this limit the incremental resistance becomes markedly dependent on pressure and field, and ultimately changes from negative to positive.

If the inner electrode is the cathode and is split, the incremental resistance is always positive. If either the outer or the inner electrode is the anode and is split the incremental resistance between the two sections is always positive. We constructed a number of tubes in which a discharge took place, without any magnetic field, between two electrodes one of which was split into two parts,



each carrying some of the current. In no case were we able to obtain a negative incremental resistance. It appears then that in split-electrode tubes like these, a negative resistance exists between electrode sections only when the split electrode is the outer electrode and is the cathode, and then only in the presence of a magnetic field. It may be that the negative resistance is associated with the electron transfer across the central plane.

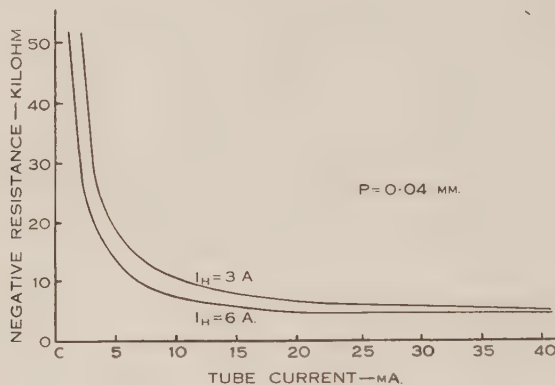


Fig. 11.—Initial incremental negative resistance as a function of total tube current.

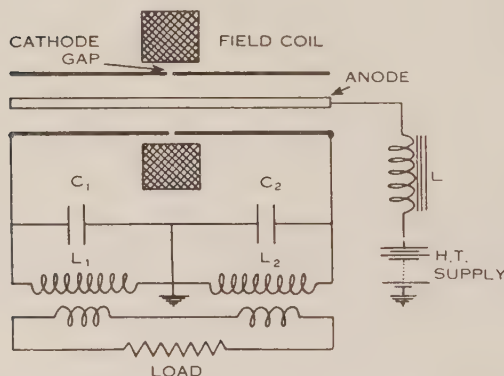


Fig. 12.—Circuit of oscillator using split-cathode discharge tube.

## V. USE OF THE DISCHARGE TO GENERATE OSCILLATIONS

Because the discharge in a split-cathode tube may have a negative resistance characteristic, as described in Section IV, it may be used to sustain oscillations in a circuit as in other "negative resistance" oscillators. A convenient circuit for such an oscillator is shown in Figure 12.

The products  $L_1 C_1$ ,  $L_2 C_2$  are adjusted to equality, although this is not a critical condition for oscillation. The choke  $L$  is not essential for the existence of oscillations but contributes to their stability.



The highest frequency we have obtained with such an oscillator is about 90 kc/sec. In the region of a few kilocycles per second the frequency is determined by the product  $L_1C_1$  ( $=L_2C_2$ ), but as the upper limit is approached the frequency becomes dependent on anode current as well and increases with it.

With anode currents above about 10 mA. the efficiency of the oscillator is about 70 per cent., increasing somewhat as anode current increases. Oscillations were obtained at low efficiency with anode currents as low as 0.1 mA.

The oscillator operates satisfactorily at pressures between 0.14 mm. and 0.02 mm. and field coil currents between 3 and 10 amperes. Within these ranges neither pressure nor field strength is at all critical. Powers in excess of 50 watts at an efficiency of about 80 per cent. were obtained with an anode current of 65 mA. and a high tension supply of 1000 volts, but such operation was possible only for short periods because of the presence of waxed joints in the tube. The maximum output for continuous rating was 8 watts at 75 per cent. efficiency with anode current 15 mA. and high tension 700 volts. The oscillator will operate from a H.T. supply as low as 300 volts with a power output of about 0.5 watt.

During oscillation the discharge appears to move to and fro through a distance of 5 cm. or more. In a tube with a longitudinal slit in the cathode, permitting visual inspection, the static discharge appears as a glow under the field coil having a narrow maximum of intensity in the central position. When oscillations begin, the glow spreads out over a much greater width with one or more maxima of intensity towards each end. Another tube was made in which the anode was divided transversely into three sections. The central section, about 1.5 cm. long, was placed symmetrically opposite the cathode gap. Under static conditions current flowed to the central anode only, but when oscillation commenced, current began to flow to the more remote anode sections until these sections carried the greater part of the total current. The current to each anode section consisted of a series of rectangular pulses, the frequency of those to the central anode section being twice that of those to the outer sections consistently with to-and-fro motion of a narrow discharge.

The life of a sealed-off oscillator of this type would be limited by clean-up of gas and under normal operating conditions would not be more than a few hours. The use of vapour such as that of mercury might lead to an oscillator with reasonable life.

## VI. ACKNOWLEDGMENTS

We wish to thank Mr. C. T. Grainger for help in taking observations and Professor V. A. Bailey and Dr. and Mrs. R. E. B. Makinson for the benefit of discussions during the course of the investigations. The work of two of us (K. S. W. Champion and E. K. Bigg) was made possible by Commonwealth Research Grants.

# TRANSLATION AND DEVELOPMENT IN TWO-DIMENSIONAL FIELDS, WITH SPECIAL REFERENCE TO PRESSURE VARIATIONS

By R. W. JAMES\*

[Manuscript received September 29, 1948]

## Summary

In previous studies of translation of field entities attention has been focused on singular points and lines of the field, with the result that the complementary concept of development in the field has been left either not at all, or quite arbitrarily, defined. By determining the least-squares fit to the rate of change of the entity *throughout the field*, rigid and unique definitions are derived of both the translation and rotation of the system of isolines of the entity, and also of the local and total development in the field.

The practical result is to derive a system of weighting factors for the translation of elements of the field, and to stress the special importance of those regions where the gradient of the entity is greatest. In meteorology this means that the movement of pressure-systems is dominated by that of the regions of strongest wind.

Comparison is made with earlier results, and the question of the consistent motion of coexistent fields (e.g. pressure and temperature) is discussed.

## I. INTRODUCTION

In many branches of physics a map of equal level lines is considered the best method of summarizing the spatial characteristics of a field distribution. A comparison of two maps drawn at different times reveals the changes that have been going on in the field during the period.

Such changes may be of two general types, a characteristic of the field may be present on both maps, but be displaced on one relative to the other, or a feature on one map may be absent on the other or present in intensified form.

A map of sea-level atmospheric pressure commonly reveals distinctive structures, centres of high or low pressure, troughs and wedges, etc., which may persist in time while moving in space, or may intensify or weaken.

The phenomenon of a bodily translation of a distinctive structure in the pressure-field is often referred to as an advection, but we prefer the term translation in what follows. A change in the structure of the field is termed a development.

A formal kinematic solution can be given of the speed of translation of a field singularity, a centre of low pressure for example, but is a singularity of such unique physical importance? What the meteorologist is fundamentally interested in is the depression itself, a wind and pressure-field covering a wide area, of which the centre of minimum pressure is an important characteristic,

\* Section of Meteorological Physics, C.S.I.R.

but no more. We assume that the depression represents a physical entity, the effects of which are felt over a wide area, and not at one point only, the translation of which must be defined in terms of what is happening throughout this area.

Once translation in a field has been defined, development is defined unambiguously as the residual change which cannot be accounted for by translation in the field.

The following discussion of translation and development will be concerned mainly with pressure-fields, for it is in meteorology that the conception of changing field-patterns is of most immediate concern.

## II. THE DISPLACEMENT OF PRESSURE-SYSTEMS

The formal geometry of changing pressure-fields has been discussed by Giao(1), and in great detail by Petterssen(2, 3).

The pressure-tendency  $\delta p/\delta t$ , as measured by an observer moving with velocity  $C$  is related to the local tendency at a fixed, momentarily coincident point,  $\partial p/\partial t$ , by the relation

$$\frac{\delta p}{\delta t} = \frac{\partial p}{\partial t} + C \cdot \nabla p \dots\dots\dots (1)$$

If the observer moves in the instantaneous direction of an isobar

$$C \cdot \nabla p = 0$$

and hence

$$\frac{\delta p}{\delta t} = \frac{\partial p}{\partial t}.$$

Thus an observer can move at an arbitrary speed along the isobars without the motion being reflected in the pressure tendency as measured by him. When he has a velocity component along the pressure-gradient,  $C \cdot \nabla p$  is not zero, and his measured tendency will differ from the local tendency.

If the observer is constrained to move so that his measured tendency is zero, his velocity will be given by

$$\frac{\delta p}{\delta t} = \frac{\partial p}{\partial t} + C \cdot \nabla p = 0.$$

The velocity  $C$  is arbitrary to the extent that an arbitrary velocity transverse to the gradient of pressure is possible. We may take out this arbitrariness by requiring that the observer move with the least possible speed consistent with his maintaining  $\delta p/\delta t = 0$ . We thus arrive at a precise definition of the speed of translation of a *point* in a pressure-field in the direction of the pressure-gradient,  $x$ ,

$$C = - \frac{\partial p}{\partial t} \bigg/ \frac{\partial p}{\partial x} \dots\dots\dots (2)$$

Note that an analysis of the Petterssen type applies to a point in the pressure-field, and gives one component of velocity only, that along the pressure-gradient. In terms of the geostrophic wind in the  $y$ -direction,  $V_g$ , we may write (2)

$$C = -\frac{\partial p}{\partial t} \bigg/ (2\rho\omega \sin \varphi V_g) \dots\dots\dots (2a)$$

where  $\rho$  is the density of air,  $\omega$  the earth's rotation, and  $\varphi$  the latitude, the motion of the system being eastward when  $\partial p/\partial t$  is negative for southerly winds in the Northern Hemisphere, and northerlies in the Southern Hemisphere.

The speed of translation of a point in any scalar field,  $A$ , say, along the gradient of  $A$  is given by an equation of the form (2) with  $A$  substituted for  $p$ .

The translation of a point in a vector-field raises new issues. Consider, for example, a pressure-gradient field  $\left(\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}\right)$ . In order to maintain his position in such a field, an observer must move with the velocity  $u, v$  in the  $x$ - and  $y$ -directions given by

$$\left. \begin{aligned} \frac{\delta}{\delta t} \left( \frac{\partial p}{\partial x} \right) &= 0 = \frac{\partial^2 p}{\partial t \partial x} + u \frac{\partial^2 p}{\partial x^2} + v \frac{\partial^2 p}{\partial x \partial y} \\ \frac{\delta}{\delta t} \left( \frac{\partial p}{\partial y} \right) &= 0 = \frac{\partial^2 p}{\partial t \partial y} + u \frac{\partial^2 p}{\partial x \partial y} + v \frac{\partial^2 p}{\partial y^2} \end{aligned} \right\} \dots\dots\dots (3)$$

These equations determine unequivocally the velocity of translation of a point in a pressure-gradient field; there is no degree of arbitrariness as is represented in a scalar field by motion along the isolines.

Solving the simultaneous equations (3) for  $u$  and  $v$ , we find

$$\left. \begin{aligned} u &= \frac{\frac{\partial^2 p}{\partial x \partial y} \cdot \frac{\partial^2 p}{\partial t \partial y} - \frac{\partial^2 p}{\partial y^2} \cdot \frac{\partial^2 p}{\partial t \partial x}}{\frac{\partial^2 p}{\partial x^2} \cdot \frac{\partial^2 p}{\partial y^2} - \left( \frac{\partial^2 p}{\partial x \partial y} \right)^2} \\ v &= \frac{\frac{\partial^2 p}{\partial x \partial y} \cdot \frac{\partial^2 p}{\partial t \partial x} - \frac{\partial^2 p}{\partial x^2} \cdot \frac{\partial^2 p}{\partial t \partial y}}{\frac{\partial^2 p}{\partial x^2} \cdot \frac{\partial^2 p}{\partial y^2} - \left( \frac{\partial^2 p}{\partial x \partial y} \right)^2} \end{aligned} \right\} \dots\dots\dots (4)$$

These equations may be simplified if the direction of translation is given. Suppose the system is translating along the  $x$ -axis with velocity  $C$ . Then  $v=0$ , i.e.

$$\frac{\partial^2 p}{\partial x \partial y} \cdot \frac{\partial^2 p}{\partial t \partial x} - \frac{\partial^2 p}{\partial x^2} \cdot \frac{\partial^2 p}{\partial t \partial y} = 0.$$

Eliminating  $\partial^2 p/\partial x \partial y$  between this expression and the first of (4) we find

$$C = u = -\frac{\partial^2 p}{\partial x \partial t} \bigg/ \frac{\partial^2 p}{\partial x^2} \dots\dots\dots (5)$$

Equation (5) gives the speed of translation of a point in a pressure-gradient field when the direction of motion is specified. In particular we may take a point on the axis of a straight trough (in the  $y$ -direction) which is translating at right angles to itself. Such a point is characterized by the fact that  $\partial p/\partial x=0$ , and the speed of translation of a point in the trough-line is given by (5).



A trough-line can best be defined as the locus of points at which the isobaric curvature is a maximum.\* When the trough-line is curved, it can no longer be characterized by the orthogonal property,  $\partial p / \partial x = 0$ . The locus of maximum curvatures is difficult to treat, but for a nearly straight trough the translation of a line orthogonal to the isobars and passing through a given point on the trough-line is an approximation to the translation of the trough-line itself in the vicinity of the point selected. Thus the expression for the speed of a straight trough is approximately valid for the speed of a limited length of curved trough.

This expression for the translation of a straight trough or wedge was given by Petterssen, but we must note that it is only valid when the pressure-field is not translating along the trough axis. It is further to be observed that the expression applies only to a point in the trough-line and not to the trough-line as a whole.

The most general uniform motion of a plane rigid body is a translation plus a rotation. Taking the speed of a trough-line at a point as characterizing the trough-line as a whole is equivalent to ignoring the possibility of a rotation in the pressure-field.

We are now in a position to discuss what is meant by the velocity of translation of a pressure-system. It is evident that the speed of translation of a point in a pressure-field along the gradient cannot be taken as the speed of translation of the system as a whole. The velocity of a point in the pressure-gradient field taken at random would seem even more unpromising as a measure of field translation, since pressure-gradient cannot be measured as accurately as pressure.

However, as Petterssen shows, the speed of translation of a field singularity offers, in many ways, a satisfactory measure of the speed of translation of the system as a whole. In some cases the more exact equations (4) would have to be taken rather than a simpler equation of the form (5), but this is no departure from the principle of taking the speed of singularities in the field as characteristic of the movement of the field as a whole.

If we define the translation of a system in these terms, the numerical value of development is specified everywhere in the field where it is possible to measure both local tendency and pressure-gradient. However, it will be appreciated that there is a good deal of arbitrariness in such a definition as the speed of a system is defined in terms of what is happening in a particular region of the field only. The evaluation of  $C$  is dictated in practice by the data available, and hence misleading results may be obtained if a well-documented area of a trough-line has not a translation typical of the trough as a whole.

It seems desirable that such a fundamental concept as development in a field of pressure should not be at the mercy of such arbitrariness. It should be possible, conceptually at least, to define the velocity of a pressure-field without ambiguity, in terms of what is observed throughout the field, and not at a point, or along a line only.

\* This definition is given in "Weather Glossary", U.S. Dep. Commerce 1946.



Fortunately it is possible to do this. In some way we must weight observations at our disposal from all parts of the field, so that the velocity of the total system is a weighted mean of the velocities of its individual parts.

### III. THE SPEED OF BEST FIT

The principle of fitting normally employed in problems such as this is the principle of least squares. We have to select a velocity  $C$ , made up of a uniform translation and a rotation, such that the residual pressure-tendency, or development

$$\frac{\delta p}{\delta t} = \frac{\partial p}{\partial t} + \mathbf{C} \cdot \nabla p$$

is as small numerically as possible everywhere in the field.

The least-squares solution is that which minimizes the integral

$$I = \int \left( \frac{\partial p}{\partial t} + \mathbf{C} \cdot \nabla p \right)^2 d\sigma \quad \dots \dots \dots (6)$$

the integral being taken over the whole area of the field of which  $d\sigma$  is an element.

In the general case we allow a rotation of the field about the origin of coordinates. The velocity of a point in the field is then

$$\mathbf{C} = \mathbf{i}(u - \Omega y) + \mathbf{j}(v + \Omega x)$$

where  $\mathbf{i}$  and  $\mathbf{j}$  are unit vectors in the  $x$ - and  $y$ -directions,  $u$ ,  $v$  are the  $x$ ,  $y$  components of the velocity of translation, and  $\Omega$  is the rotation of the field. The integral (6) may then be written

$$I = \int \left( \frac{\partial p}{\partial t} + (u - \Omega y) \cdot \frac{\partial p}{\partial x} + (v + \Omega x) \frac{\partial p}{\partial y} \right)^2 d\sigma.$$

The integral is a minimum when

$$\frac{\partial I}{\partial u} = \frac{\partial I}{\partial v} = \frac{\partial I}{\partial \Omega} = 0.$$

These conditions give the three normal equations

$$\left. \begin{aligned} \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial x} \cdot d\sigma + \int \frac{\partial p}{\partial x} \left\{ (u - \Omega y) \frac{\partial p}{\partial x} + (v + \Omega x) \frac{\partial p}{\partial y} \right\} d\sigma &= 0 \\ \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial y} \cdot d\sigma + \int \frac{\partial p}{\partial y} \left\{ (u - \Omega y) \frac{\partial p}{\partial x} + (v + \Omega x) \frac{\partial p}{\partial y} \right\} d\sigma &= 0 \\ \int \frac{\partial p}{\partial t} \left\{ -y \frac{\partial p}{\partial x} + x \frac{\partial p}{\partial y} \right\} d\sigma + \int \left[ -y \frac{\partial p}{\partial x} + x \frac{\partial p}{\partial y} \right] \left\{ (u - \Omega y) \frac{\partial p}{\partial x} + (v + \Omega x) \frac{\partial p}{\partial y} \right\} d\sigma &= 0 \end{aligned} \right\} \dots \dots \dots (7)$$

A rotation  $\Omega$ , about the point  $x_0$ ,  $y_0$  would give exactly the same solution, except that  $u$  would have to be replaced by  $u + \Omega y_0$  and  $v$  by  $v - \Omega x_0$ . The point of rotation in the field is therefore arbitrary, a shift in the centre of rotation being brought about by the addition of a term to the translation.

When there is no rotation the solution of (7) is

$$\left. \begin{aligned} u &= - \frac{\int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial x} d\sigma \cdot \int \left(\frac{\partial p}{\partial y}\right)^2 d\sigma - \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial y} d\sigma \cdot \int \frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} d\sigma}{\int \left(\frac{\partial p}{\partial x}\right)^2 d\sigma \cdot \int \left(\frac{\partial p}{\partial y}\right)^2 d\sigma - \left\{ \int \frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} d\sigma \right\}^2} \\ v &= - \frac{\int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial y} d\sigma \cdot \int \left(\frac{\partial p}{\partial x}\right)^2 d\sigma - \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial x} d\sigma \cdot \int \frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} d\sigma}{\int \left(\frac{\partial p}{\partial x}\right)^2 d\sigma \cdot \int \left(\frac{\partial p}{\partial y}\right)^2 d\sigma - \left\{ \int \frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} d\sigma \right\}^2} \end{aligned} \right\} \dots (8)$$

Equations (8) give the velocity of translation of a pressure-system as a whole where there is no rotation in the field in terms of the pressure-tendency and pressure-gradient everywhere in the field. It simplifies greatly for a field symmetrical about a point or a line. Consider, for example, the points  $A$  ( $x, y$ ),  $B$  ( $-x, y$ ) about an axis of symmetry, taken as the  $y$ -axis. We have

$$\left(\frac{\partial p}{\partial y}\right)_A = \left(\frac{\partial p}{\partial y}\right)_B ; \quad \left(\frac{\partial p}{\partial x}\right)_A = -\left(\frac{\partial p}{\partial x}\right)_B.$$

In such a system the contributions  $\frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} \cdot d\sigma$  to the integral  $\int \frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} \cdot d\sigma$  cancel in pairs, so that the whole integral is zero.

Under these conditions equations (8) become

$$\left. \begin{aligned} u &= - \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial x} d\sigma / \int \left(\frac{\partial p}{\partial x}\right)^2 d\sigma \\ v &= - \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial y} d\sigma / \int \left(\frac{\partial p}{\partial y}\right)^2 d\sigma \end{aligned} \right\} \dots (9)$$

Equations (8) also simplify when the direction of the field-translation is specified. If we take the motion to be in the  $x$ -direction we have  $v=0$ , i.e.

$$\frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial y} d\sigma \cdot \int \left(\frac{\partial p}{\partial x}\right)^2 d\sigma - \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial x} d\sigma \cdot \int \frac{\partial p}{\partial x} \cdot \frac{\partial p}{\partial y} d\sigma.$$

Eliminating  $\int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial y} d\sigma$  between this expression and the first of equations

(8), we find for the speed of translation of the system

$$C = U = - \int \frac{\partial p}{\partial t} \cdot \frac{\partial p}{\partial x} d\sigma / \int \left(\frac{\partial p}{\partial x}\right)^2 d\sigma \dots (10)$$

This result is perfectly general, and applies to asymmetrical pressure-systems.

Equation (10) is analogous to the Petterssen(3, p. 385) expression for the movement of a point in an isobar along the gradient of pressure,

$$C = - \frac{\partial p}{\partial t} / \frac{\partial p}{\partial x}.$$

The difference is that the numerator and denominator in this simple expression are weighted with the pressure-gradient at each point, and summed separately. Pressure-tendencies throughout the field enter into the determination of  $C$ , but most weight is given to those parts of the field where the pressure-gradient in the direction of motion is tightest, and where equation (10) cannot be evaluated for the entire field the best indication of the translation of the system is given by the speed of the isobars where the gradient is tightest in the direction of motion.

This contrasts with the Petterssen analysis, which assigns all the weight to tendencies at singular points, where the pressure-gradient in the direction of motion is zero. As regards the practical application of the two methods, all that need be said is that second derivatives are normally more difficult to evaluate than simple derivatives and from this point of view the movement of the isobars will generally be a more trustworthy measure of translation than the movement of pressure-gradients.

#### IV. THE MEANING OF DEVELOPMENT

Development is normally conceived as that part of the pressure-change which remains when the translation (and rotation) of the field is allowed for. It can only have physical meaning when the translation refers to the field as a whole rather than to some particular point of it. An unambiguous definition of development requires a unique definition of translation, and this is afforded by the least-squares criterion. We accordingly define a development in a field of pressure as that part of a pressure-change which remains when a least-squares fit to the translation and rotation of the field has been taken out.

It is convenient to distinguish between a local development, which is the residual change of pressure at a point, and overall development in a field of pressure, which is the integrated effect of local developments throughout the field.

We might also distinguish between overall development proper and a distortion in the field. For example, if a depression were elongating, the fall of pressure along the semi-major axis would be compensated by a rise along the semi-minor axis, and the net development in the field would be zero. If we take the entire earth as our pressure-field, development in one region must always be compensated elsewhere, for the total pressure must remain constant, so from this viewpoint all development may be regarded as distortion. The conception of development thus depends on our specifying a delimited and identifiable pressure-pattern, overall development representing an exchange of mass with the region outside the system, and a distortion representing a rearrangement within the system.

Local development has been defined as a residual rate of change of pressure at a point and is therefore a strictly quantitative concept. The overall development, which is the integral of this, is also a quantitative concept, and it is convenient to coin a name for this quantity. We accordingly define the *strengthening\** of a pressure-system, with certain conventions outlined below, as

\* The need for a generally improved terminology has been discussed elsewhere.

$$\int \frac{\delta p}{\delta t} d\sigma = \int \frac{\partial p}{\partial t} d\sigma + \int \mathbf{C} \cdot \nabla p d\sigma.$$

The second integral is zero if we take a system contained in a closed isobar. It is likewise zero if the area of integration is large enough to contain the entire translating configuration, but not so large as to embrace extraneous pressure-systems. Under these conditions we may write

$$\int \frac{\delta p}{\delta t} d\sigma = \frac{\partial}{\partial t} \int p d\sigma.$$

We say that the system is strengthening if the existing contrast of pressure is being intensified; for a cyclone if overall pressure is low and falling still further. Weakening is applied to the converse case of pressure low, but rising.

For an anticyclone strengthening denotes overall pressure higher than normal and rising, weakening, pressure high, but falling.

This conception of strengthening is a useful one, in that it suggests that cyclones, for example, can be characterized by the term *strength* which indicates the time integral of all past strengthening. We can characterize a cyclone as being strong or weak according to whether the net strengthening over past time has been great or small. Since strengthening is the time rate of removal of air from the area of low pressure, the strength of a low is defined by the total removal of air from a uniform field of pressure required to produce the low. Thus we can characterize a depression quantitatively in terms of the total withdrawal of air required to produce it.

Similarly a high may be characterized by the total weight of air added to a uniform pressure-field to produce it.

## V. TRANSLATION OF A THERMAL-FIELD

So far attention has been confined to translation and development in a field of pressure. Exactly the same considerations apply to a field of temperature. Given temperature and isallotherm fields, the least-squares determination of the speed of translation is

$$C = - \int \frac{\partial T}{\partial t} \cdot \frac{\partial T}{\partial x} \cdot d\sigma / \int \left( \frac{\partial T}{\partial x} \right)^2 d\sigma \dots\dots\dots (11)$$

where  $x$  is the direction in which the field is translating. If the direction of motion is not known we should have to use a more complex expression analogous to equations (8) for pressure translation.

Having defined the speed of translation of a thermal-field it is possible to define a local temperature development by

$$\frac{\delta T}{\delta t} = \frac{\partial T}{\partial t} + \mathbf{C} \cdot \nabla T \dots\dots\dots (12)$$

and an overall development.



Overall development is not, however, of such practical importance in a thermal-field as distortion, which may take the form of a tightening or relaxing of the temperature-gradient in some particular region of the field.

The most important case of thermal-field translation is the movement of a front, which is a line of discontinuity, or of rapid transition of temperature. Equation (11) applies to the translation of a zone of temperature transition, and, in the limit to a true discontinuity. It is to be remembered, however, that fronts normally do not move with the same speed throughout their length, and in applying a least-squares fit to the translation of a frontal zone, it may be necessary to allow for a rotation as well as a translation.

It will be noted that the part of the field in which the temperature-gradient is greatest in the direction of motion contributes most to the determination of the speed of translation in a least-squares fit. A single point analysis of translation should therefore be applied to a frontal region. It is thus possible to arrive at an estimate of the speed of translation of a front at right angles to itself from the expression

$$C = - \frac{\partial T}{\partial t} \bigg/ \frac{\partial T}{\partial x} \dots\dots\dots (13)$$

where  $x$  is in the direction of the temperature-gradient, i.e. at right angles to the front.

Assuming that winds are approximately geostrophic, we may express  $\partial T/\partial x$  in terms of the change with height of wind speed in the direction of the front from the thermal-wind relation

$$C \approx -g \frac{\partial T}{\partial t} \bigg/ \left( 2\omega \sin \phi T \frac{\partial v}{\partial z} \right) \dots\dots\dots (13a)$$

In numerical terms, for latitude  $45^\circ$ ,  $T=300^\circ \text{ K.}$ , when the tendency is expressed in  $^\circ \text{F. per 3 hr.}$ , shear in m.p.h. per 1000 ft., we have

$$C \text{ (m.p.h.)} = -25 \frac{\partial T}{\partial t} \bigg/ \frac{\partial v}{\partial z}.$$

Thus a temperature fall of  $5^\circ \text{ F.}$  in 3 hours associated with a southerly wind increase of 5 m.p.h. per 1000 ft. indicates a translation speed of 25 m.p.h. towards the east in the Southern Hemisphere.

This expression for the speed of a diffuse front at right angles to itself is not always easy to apply, especially in the surface layers, where temperature is notoriously unrepresentative. The difficulty in applying it to upper air data is that raysonde ascents are rarely released at sufficiently frequent intervals to permit an estimation of  $\partial T/\partial t$ .

A sharp temperature discontinuity is associated with a discontinuity in the pressure-gradient, and there is no reason why a front should not be defined by this pressure-gradient discontinuity, or sharp change. This, Petterssen does in his expression for the speed of a sharp front, which gives the speed of translation of the wind-shear across the front. Where the temperature transition zone



is broad, the frontal discontinuity in pressure-gradient is replaced by a trough, and the speed of the front has to be taken as the speed of the trough.

The important point here is that the speed of a front may be estimated in two quite different ways, either from the speed of the thermal-field, or the speed of the attendant pressure-field.

Treating a front as a thermal-field is of theoretical value, in that it makes clear the meaning of frontogenesis. The distortion of a temperature field mentioned earlier, in which temperature-gradient is tightened or relaxed in some area of the field is what is commonly implied by the terms frontogenesis and frontolysis.

Ideally, a least-squares analysis of the temperature-field will give a uniform translation and a rotation of the field, and a field of local development. Examination of this development field would indicate the regions of frontogenesis or frontolysis.

## VI. INVARIANT TRANSLATION SYSTEMS

By least-squares fitting it is possible to assign a speed of translation and a rotation to all meteorological field quantities. For example, the speed of translation in the  $x$ -direction of a field quantity  $A$ , which may be pressure, temperature, density, specific humidity, wind velocity, etc. or derivatives of these of any order is

$$C = - \int \frac{\partial A}{\partial t} \cdot \frac{\partial A}{\partial x} d\sigma / \int \left( \frac{\partial A}{\partial x} \right)^2 d\sigma$$

where  $A$  is of the general form

$$A = \frac{\partial^{m+n+r} \psi}{\partial x^m \partial y^n \partial t^r}$$

where

$$\psi = (p, T, \rho, V).$$

We term an invariant translating system one in which a least-squares fit to every conceivable field quantity gives the same speed and direction of translation. Only with an invariant system do all the parts of the various field quantities bear permanently the same relation to each other. The higher derivatives of meteorological quantities are not sufficiently accurately known for us to specify an invariant system in the above rigid sense, but the conception is useful as applied to scalar field quantities and their first derivatives only.

It will be noticed that invariance (or its absence) is a different conception from development, which has been taken to embrace the relations within a single field only. It is conceivable—mathematical analysis is required to prove whether or not it is possible—that a pressure-system could translate unaltered through a stationary thermal-field. Here there is no development either in the pressure or the temperature-fields, but the system is not invariant in the sense that the pressure-field is changing in relation to the thermal-field.

Development, as defined above, is a conception applying essentially to a horizontal section of the atmosphere. Invariance might be applied, not only to the relations of two fields in the same horizontal plane, but to the same field at different heights. Thus if the vertical axis of a low is tilting, the surface-pressure-

field and also the pressure-field at height may remain unchanged in motion but the speeds of the system at the two levels will not be the same.

We may term such a phenomenon inconsistent translation.

The most important application of the idea of an inconsistent translation is to fronts. We have seen that a front has a dual character of a temperature discontinuity or rapid transition, and a pressure discontinuity or rapid variation. If a front is to preserve its identity in movement, both features must remain together, so that the velocity of a front derived from its aspect as a thermal-field must agree with that deduced from its role as a pressure-field.

#### VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Section of Meteorological Physics, C.S.I.R. The author is indebted to Mr. C. H. B. Priestley, Officer-in-Charge of the Section, for helpful criticism.

#### VIII. REFERENCES

- (1) GIAO, A.—*Mémor. Off. Nat. Mét. Fr.* **20**: 41 (1929).
- (2) PETERSSEN, S.—*Geofys. Publ.* **10** (2): 1-92 (1933).
- (3) PETERSSEN, S.—“Weather Analysis and Forecasting”, pp. 378-440. (McGraw-Hill Book Co.: New York and London, 1940.)

# THE INFLUENCE OF VERTICAL IONIC DRIFT ON A "CHAPMAN REGION"

By C. B. KIRKPATRICK\*

[Manuscript received September 3, 1948]

## Summary

Vertical transport of ions is taken into account in a theoretical treatment of an ionized region in which ion production and decay are operative. Methods are evolved for computing the electron density in both a "moving" and a static layer. Ionic drift is believed to be caused by tidal motion of the upper air and so, for conditions representative of the  $E$  region and the solar atmospheric tide, the effect of ionic drift is investigated and the theoretical departure from normal "Chapman" behaviour is measured.

## I. INTRODUCTION

A comprehensive theory of the formation of ionized regions in the atmosphere has been developed by Chapman(1). He showed that the absorption of monochromatic solar radiation by an ionizable atmospheric gas produces ionization at the rate

$$I = I_0 e^{(1 - z - e^{-z} \sec \chi)}$$

where  $z$  is the reduced height (i.e. height measured, above an arbitrary datum level, in units of the scale height  $H$ ),  $\chi$  the solar angle, and  $I_0$  proportional to the absorption coefficient of the atmospheric constituent.

In general, free electrons may disappear by a complex series of processes, simple radiative recombination, three-body recombination, and attachment and detachment(2). Nevertheless, Massey and his collaborators have shown, in a series of papers, that the overall result of these processes is to cause the disappearance of free electrons according to a simple recombination law, so that, neglecting diffusion, the electron density  $N$  is given by

$$\frac{dN}{dt} = I - \alpha N^2$$

where  $\alpha$  is the "effective" recombination coefficient.

Chapman showed that, under such conditions, a pronounced maximum in  $N$  occurs at a certain level in the atmosphere. In other words an ionized region, now commonly called a "Chapman region" is formed. The properties of such regions have been studied in detail by Chapman, and by Millington(3, 4). They agree substantially with those of the  $E$  and  $F_1$  regions of the ionosphere, as elucidated by Appleton and his co-workers.

\* Lecturer in Mathematics, Sydney Technical College.

Nevertheless, the examination of data from certain ionospheric observatories over a period of ten years has revealed that the  $E$  and  $F_1$  regions show small but regular departures from "Chapman" behaviour (Martyn, personal communication). Martyn(5) has advanced a theory that the known large departures of the  $F_2$  region from "Chapman" characteristics are caused by the action of the solar atmospheric tide. According to this theory, tidal motion of the air, though predominantly horizontal, gives rise, in the presence of the earth's magnetic field, to vertical ionic drift, electrons and ions moving in the same sense. Martyn has investigated the effect of such drift on the  $F_2$  region at night, when ion production and decay may be neglected.

It seems likely that the observed small anomalies in the  $E$  and  $F_1$  regions may also be produced in this manner. This hypothesis can only be tested for day-time conditions, however, since these regions disappear at night. Hence it is necessary to take account of both ionization production and decay in solving the relevant equations. This introduces complexities whose solution is attempted in this paper. This solution would facilitate the interpretation of the anomalies found in the  $E$  and  $F_1$  regions in terms of tidal movements, and should permit the identification of the phase of the solar tide in these regions. In turn, this identification should throw considerable light on the problem of the location of the ionospheric currents producing the solar diurnal magnetic variations.

## II. ELECTRON DISTRIBUTION IN A STATIC LAYER

Consider a region of the atmosphere which is static relative to the earth. If  $I$  is the rate of ion production and free electrons disappear as if by a recombination law, then  $N$ , the electron number density at height  $z$  at time  $t$ , obeys the law

$$\frac{dN}{dt} = I(z, t) - \alpha N^2 \quad \dots \dots \dots (1)$$

where  $\alpha$  is the effective recombination coefficient and will be assumed constant in the region considered.

During the night, if it is assumed that no ion production occurs,

$$\frac{dN}{dt} = -\alpha N^2 \quad \dots \dots \dots (2)$$

so that

$$\frac{1}{N} = \frac{1}{N_s} + \alpha(t - t_s) \quad \dots \dots \dots (3)$$

the subscript  $s$  denoting sunset values.

Equation (1) may be written

$$\frac{dN}{dt} = \alpha(\eta^2 - N^2)$$

where

$$\eta = \sqrt{I/\alpha} \quad \dots \dots \dots (4)$$

Usually it is assumed that  $\frac{dN}{dt}$  is small, except near sunrise and sunset; then

$\eta$  has physical significance as an approximation to  $N$ . The above assumption, however, will not be made here.

$$\text{Let} \quad I = I_0 [\sigma(z, t)]^2, \quad \dots\dots\dots (5) \\ 0 \leq \sigma \leq 1;$$

$$\eta_0 = \sqrt{I_0/\alpha} \quad \dots\dots\dots (6)$$

$$\text{Then} \quad \eta = \eta_0 \cdot \sigma(z, t) \quad \dots\dots\dots (7)$$

Let  $n$  denote electron density expressed in units of  $\eta_0$ .

$$\text{Then} \quad N = \eta_0 \cdot n(z, t) \quad \dots\dots\dots (8)$$

and equation (1) becomes

$$dn/dt = \alpha \eta_0 \cdot (\sigma^2 - n^2) \quad \dots\dots\dots (9)$$

It will be convenient to express time by angular measure  $\omega t$  and to introduce a variable  $\theta$  such that

$$d\theta = d(\omega t).$$

Equation (9) may now be written

$$\frac{dn}{d\theta} = K \left( \frac{\sigma^2 - n^2}{2} \right) \quad \dots\dots\dots (10)$$

where

$$K = \frac{2\alpha\eta_0}{\omega} \quad \dots\dots\dots (11)$$

Manipulation of the right-hand side of (10) gives

$$K(\sigma n - n^2) + K \frac{(\sigma - n)^2}{2} = \frac{dn}{d\theta} = K(\sigma^2 - \sigma n) - K \frac{(\sigma - n)^2}{2} \quad \dots\dots\dots (12)$$

and so

$$K(\sigma n - n^2) \leq dn/d\theta \leq K(\sigma^2 - \sigma n) \quad \dots\dots\dots (13)$$

It follows that

$$\frac{dn}{d\theta} + K\sigma \cdot n \leq K\sigma^2 \quad \dots\dots\dots (14)$$

$$\frac{d}{d\theta} \left( \frac{1}{n} \right) + K\sigma \cdot \left( \frac{1}{n} \right) \leq K \quad \dots\dots\dots (15)$$

Let subscripts 0, 1 denote values at times  $t_0$ ,  $t_1$  respectively and let  $t_0 \leq t \leq t_1$ . Introducing the integrating factor  $e^{-K\alpha}$  where

$$a = a(\theta) = \int_{\theta}^{\theta_1} \sigma(\theta) \cdot d\theta, \quad \dots\dots\dots (16)$$

and integrating between the terminal values  $\theta_0$  and  $\theta_1$ , equations (14) and (15) yield respectively

$$n_1 \leq n_0 e^{-KA} + \int_{e^{-KA}}^1 \frac{\sigma}{e^{-KA}} \cdot d[e^{-KA}] = p_1 \quad \dots\dots\dots (17)$$

$$\frac{1}{n_1} \leq \frac{1}{n_0} \cdot e^{-KA} + \int_{e^{-KA}}^1 \frac{1}{\frac{\sigma}{e^{-KA}}} \cdot d[e^{-KA}] = \frac{1}{q_1} \quad \dots\dots\dots (18)$$

where  $A = a(\theta_0)$  and  $z = \text{constant}$ .



Hence, discarding subscripts,  $q < n < p$ , and  $n$  may be found if  $p$  and  $q$  approach equality.

Consider the expressions for  $p$  and  $q$  when

$$e^{-KA} < n_0 < 1 \dots\dots\dots (19)$$

with  $e^{-KA}$  sufficiently small to be neglected, so that

$$p = \int_0^1 \sigma \cdot d[e^{-Ka}] \dots\dots\dots (20)$$

$$\frac{1}{q} = \int_0^1 \frac{1}{\sigma} \cdot d[e^{-Ka}] = K \int_{\theta_1}^{\theta_2} e^{-Ka} \cdot d\theta \dots\dots\dots (21)$$

Substituting, in equations (20) and (21),

$$\sigma = p + \Delta\sigma$$

and imposing the condition

$$|\Delta\sigma| < p \dots\dots\dots (22)$$

it is easy to show that

$$q \simeq p \left\{ 1 - \int_0^1 \left( \frac{\Delta\sigma}{p} \right)^2 \cdot d[e^{-Ka}] \right\} \simeq p.$$

These conditions are sufficient to establish the approximations

$$n \simeq \sigma \text{ (first order)} \dots\dots\dots (23)$$

$$q \simeq n \simeq p \text{ (second order)} \dots\dots\dots (24)$$

When  $\sigma$  and  $K$  are sufficiently large for  $e^{-Ka}$  to decay rapidly from unity as time retrogresses from  $t_1$ , the time interval for the effective range of integration in (20) and (21) is curtailed and the condition (22) is not so stringent. This condition tends to be satisfied and the difference between  $p$  and  $q$  tends to diminish as  $\sigma$  approaches a sufficiently large maximum. Moreover, calculations

TABLE 1

Time of Sunrise ( $t_0$ ) .. 0600				
Height .. .. $z = 0.5$				
Initial Electron Density.. $n_0 = 0.014$				
Layer Constant .. .. $K = 0.35$				
Time	$\sigma$	$q$	$n$	$p$
0700	0.398	0.055	0.200	0.247
0800	0.700	0.647	0.667	0.667
0830	0.780	0.762	0.762	0.762
1600	0.700	0.723	—	0.724
1700	0.398	0.472	0.473	0.476
1800	0.000	0.123	0.166	0.277

indicate that condition (22) is not always a necessary one, the discrepancy between  $p$  and  $q$  usually being negligible except near sunrise and sunset.

A quantitative idea of the discrepancy between  $p$  and  $q$  is given in Table 1, in which some values of  $p$  and  $q$  are compared with values of  $n$ , the latter computed by numerical integration. The values refer to a typical height of a "Chapman region", later referred to as Region  $S$ . The electron density at sunrise ( $n_0$ ) was computed from the sunset value, which is independent of  $n_0$ .

It can be seen that  $p$  and  $q$  approach equality about two hours after sunrise and start to diverge appreciably about one hour before sunset. In addition,  $n$  lies closer to  $p$  near sunrise and closer to  $q$  near sunset (features which are explicable when the terms in (12) containing  $(\sigma - n)^2$  are taken into account).

The solution  $n = p$  is valid between about 8 a.m. and 4 p.m. and (20) is convenient for purposes of computation. This solution was therefore adopted. Thus, for a given  $z$ ,  $n_1 = n(z, t_1)$  is given by the area under the  $(\sigma, e^{-Ka})$  curve. Paired values of  $\sigma$  and " $a$ " may be obtained from the appropriate  $z$  contour of  $\sigma$  plotted against  $\omega t$ , " $a$ " being equal to the area under the  $\sigma$  curve between  $\omega t$  and  $\omega t_1$ .

The solution  $n = q$  is useful for obtaining  $1/N$  in a form similar to the night-time solution (3), viz.,

$$\frac{1}{N_1} = 2\alpha \int_{t_0}^{t_1} e^{-Ka} \cdot dt, \dots\dots\dots (25)$$

which illustrates the increase in electron density as ion production counteracts recombination.

### III. ELECTRON DISTRIBUTION IN A "MOVING" LAYER

#### (a) General Discussion

When ions of all kinds are subject to vertical transport (e.g. as a result of tidal motion) the electron density  $N$  satisfies the equation of continuity

$$\frac{\partial N}{\partial t} = I - \alpha N^2 + \frac{\partial(Nv)}{\partial z} \dots\dots\dots (26)$$

where  $v = v(z, t)$  is the vertical drift velocity, measured positive downwards. The motion of the ions is prescribed by the velocity  $v$ , and so the position of the cluster of ions which was at  $z_0$  at time  $t_0$  is found by solving the subsidiary equation

$$dz/dt = -v(z, t) \dots\dots\dots (27)$$

Then the variation of electron density in this cluster is found from the equation of continuity in the form

$$\frac{dN}{dt} = I - \alpha N^2 + N \cdot \frac{\partial v}{\partial z} \dots\dots\dots (28)$$

where  $d/dt$  denotes differentiation following the motion of the fluid (i.e. the ion cluster or, more accurately, an infinitely thin stratum of ionization). In the subsidiary equation (28)  $N$ ,  $I$ , and  $v$  apply to the fluid element at  $z_0$  at time  $t_0$  and are therefore functions of  $z_0$  and  $t$ . The term  $N \frac{\partial v}{\partial z}$  represents the electron

flux rate into the moving cluster of ions and should be of minor importance in regions where the velocity gradient is small.

(b) *Solution of the Subsidiary Equations*

If the ionic drift is due to tidal action  $v$  is oscillatory and we shall therefore consider the case of a progressive velocity wave

$$v = v_0 \sin x, \dots\dots\dots (29)$$

$$x = \omega t + \beta z \dots\dots\dots (29a)$$

Let

$$k = \beta v_0 / \omega \dots\dots\dots (30)$$

Then, from (27),

$$d(\omega t) = \frac{d(\beta z)}{-k \sin x} = \frac{dx}{1 - k \sin x}$$

Integrating,

$$\omega t = \int \frac{dx}{(1 - k \sin x)} = \theta + \text{constant},$$

where, without loss of generality,  $\theta$  denotes a continuous single-valued function of  $x$ , defined by the equations

$$\theta(x) = -\frac{2}{\sqrt{1-k^2}} \arctan \left\{ \frac{k - \tan \frac{x}{2}}{\sqrt{1-k^2}} \right\} \dots\dots\dots (31)$$

$$\theta(0) = -\frac{2}{\sqrt{1-k^2}} \text{Arctan} \left\{ \frac{k}{\sqrt{1-k^2}} \right\} \dots\dots\dots (31a)$$

where "Arctan" denotes the principal value.

If  $\lambda$  denotes the constant of integration,

$$\omega t - \theta = \lambda \dots\dots\dots (32)$$

and

$$\beta z + \lambda = x - \theta \dots\dots\dots (33)$$

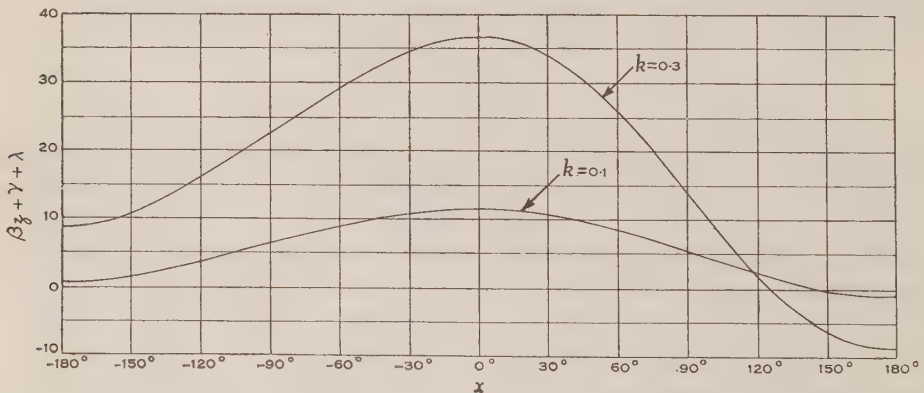


Fig. 1

Either of these equations is an implicit expression of the relation between  $z$  and  $t$  for the fluid element specified by the value assigned to the parameter  $\lambda$ . In other words the fluid element at say  $(z_0, t_0)$  may be located at any other time

by associating it with the value of  $\lambda$  obtained by putting  $z=z_0$ ,  $t=t_0$  in (32). Corresponding values of  $\beta z$ ,  $x$ , and  $\omega t$  may be found by using (33) to plot a curve of  $\beta z + \lambda$  against  $x$  as in Figure 1.

Let  $\lambda_0$  denote the value of  $\lambda$  for the fluid element at the height  $z=0$  at time  $t=0$ . Then substitution in (33) gives

$$\lambda_0 = \frac{2}{\sqrt{1-k^2}} \operatorname{Arctan} \left\{ \frac{k}{\sqrt{1-k^2}} \right\} \dots \dots \dots (34)$$

If the fluid element  $\lambda$  attains a maximum height  $\zeta$  at time  $\tau$  then it is easily seen that

$$\omega\tau = 2\pi m + \lambda - (\lambda_0 - m\delta) \dots \dots \dots (35)$$

$$\beta\zeta = (\lambda_0 - m\delta) - \lambda \dots \dots \dots (36)$$

where  $m$  is 0 or any integer and

$$\delta = 2\pi[(1-k^2)^{-\frac{1}{2}} - 1] \approx \pi k^2 \dots \dots \dots (37)$$

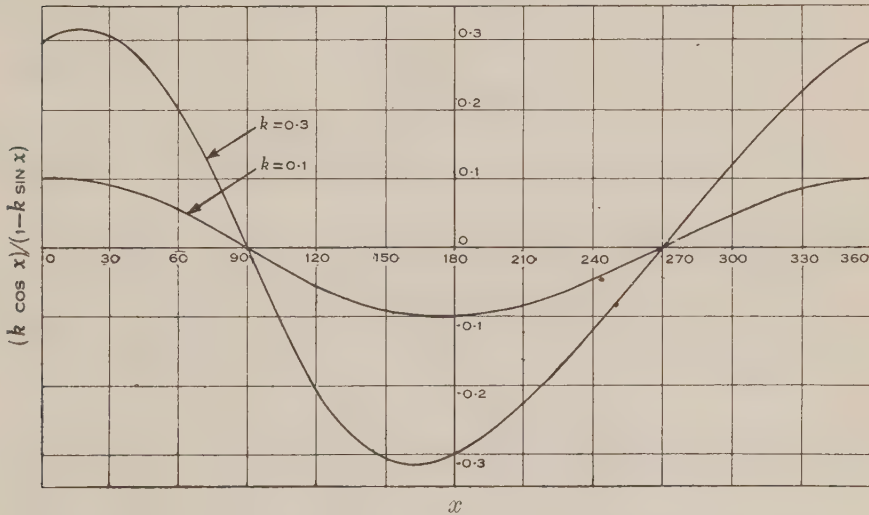


Fig. 2.—One cycle of the function  $(k \cos x)(1 - k \sin x)$  which occurs in the night-time solution. Maxima and minima have magnitude  $k/\sqrt{1-k^2}$  and occur at  $x = \arcsin k$ .

When  $m=0$ ,

$$\omega\tau = \lambda - \lambda_0$$

$$\beta\zeta = \lambda_0 - \lambda$$

and so  $\zeta \leq 0$  and  $\tau \geq 0$  according as  $\lambda \geq \lambda_0$  (see Fig. 3). It is clear that the constant  $\lambda_0 - \lambda$  may be interpreted as a maximum height (in angular measure) of the fluid element  $\lambda$ , while  $\delta$  represents the descent of each fluid element between successive maxima.

The subsidiary equation (28) may now be solved for the condition  $\lambda = \text{constant}$ . Using the notation previously introduced (28) becomes



$$\frac{dn}{d\theta} - k \cos x \cdot n = K \left\{ \frac{\sigma^2 - n^2}{2} \right\} \dots\dots\dots (38)$$

where  $\sigma = \sigma(\lambda, \theta)$  and  $\lambda = \text{constant}$ .

Writing

$$\begin{aligned} n' &= (1 - k \sin x) \cdot n \\ \sigma' &= (1 - k \sin x) \cdot \sigma \end{aligned}$$

and using the relation

$$dx/d\theta = 1 - k \sin x \dots\dots\dots (39)$$

equation (38) takes the form

$$\frac{dn'}{d\theta} = \frac{K}{2} (\sigma\sigma' - nn') \dots\dots\dots (40)$$

Hence

$$K(\sigma - n)n' \leq dn'/d\theta \leq K\sigma(\sigma' - n') \dots\dots\dots (41)$$

If  $\sigma(\lambda, \theta)$  is sufficiently large,  $\frac{\partial \sigma}{\partial \theta}$  is not too large, and (19) is satisfied,  $n \approx \sigma$  and application of the procedure adopted for solving (10) leads to the second order approximation

$$n_1' = \int_0^1 \sigma' \cdot d[e^{-Ka}] \dots\dots\dots (42)$$

where  $\lambda = \text{constant}$ , "a" is the area between  $\theta$  and  $\theta_1$  under the  $\lambda$  contour of  $\sigma$ , and  $n_1'$  is the value of  $n'$  at time  $t_1$  and a height  $z_1$  determined by the values of  $\lambda$  and  $t_1$ .

The  $q$ -type solution corresponding to (25) is

$$\frac{1}{N_1'} = 2\alpha \int_{t_0}^{t_1} \frac{e^{-Ka}}{1 - k \sin x} \cdot dt \dots\dots\dots (43)$$

### (c) Night-Time Solution

When no ion production occurs ( $I=0$ ),

$$\frac{dN}{d\theta} - k \cos x \cdot N = -\frac{\alpha}{\omega} \cdot N^2.$$

Solving, with the aid of (39),

$$\begin{aligned} \frac{1}{N_1'} &= \frac{1}{N_s'} + \frac{\alpha}{\omega} \int_{x_s}^{x_1} \frac{dx}{(1 - k \sin x)^2} \\ &= \frac{1}{N_s'} + \frac{1}{1 - k^2} \cdot \alpha \left\{ (t_1 - t_s) - \left[ \frac{k \cos x}{\omega(1 - k \sin x)} \right]_{x_s}^{x_1} \right\} \dots\dots\dots (44) \end{aligned}$$

where subscript  $s$  denotes sunset values.

For comparison with (43), the solution may be written in the form

$$\frac{1}{N_1'} = \frac{1}{N_s'} + \alpha \int_{t_s}^{t_1} \frac{dt}{(1 - k \sin x)} \dots\dots\dots (45)$$

### (d) Phase of the Velocity Function

The above solutions are applicable to the more general case

$$v = v_0 \sin(\omega t + \beta z + \gamma)$$

provided (29a) and (33) are replaced by the relations

$$x = \omega t + \beta z + \gamma \dots\dots\dots (46)$$

$$\beta z + \gamma + \lambda = x - \theta \dots\dots\dots (47)$$

### (e) Velocity Function Independent of Height

Suppose  $v$  is given by

$$v = v_0 \sin(\omega t + \gamma) \dots\dots\dots (48)$$

Then

$$dz = -v_0 \sin(\omega t + \gamma) \cdot dt,$$

$$z = z_0 - \frac{v_0}{\omega} [\cos \gamma - \cos(\omega t + \gamma)] \dots\dots\dots (49)$$

Consequently the solution (42) is replaced by

$$n_1 = n(z_0, t_1) = \int_0^1 \sigma(z_0, t) \cdot d[e^{-Kz}] \dots\dots\dots (50)$$

where  $z_0 = \text{constant}$ , and  $n_1$  is the value of  $n$  at time  $t_1$  and at the height

$$z_1 = z_0 - \frac{v_0}{\omega} [\cos \gamma - \cos(\omega t_1 + \gamma)].$$

### (f) Comparison of the Solutions

Solutions (42) and (50) show that when ion drift occurs the  $z$  contours of  $\sigma$  plotted against  $\omega t$  are replaced by other contours (e.g.  $\lambda$  contours), which account for the variation of ion production with both time and height as the fluid elements pursue their courses. In the absence of velocity-height gradient ( $\beta = 0$ ) the solution obtained for a static layer is not otherwise modified and so the layer will not be much affected when  $v$  is small. In the presence of a velocity gradient

$$\partial v / \partial z = \beta v_0 \cdot \cos x, \quad (\beta \text{ in radians})$$

the factor  $(1 - k \sin x)$  is introduced into the solution. This factor takes account of ions flowing into or out of a moving fluid element and it predominates when  $v$  is small, for the velocity gradient is then large.

If  $\beta$  is small the progressive velocity wave has a wavelength large in comparison with the thickness of the layer and the velocity gradient is, at any particular time, approximately constant throughout the layer. In such a case the layer as a whole is subject to condensation or rarefaction according as  $\partial v / \partial z$  is positive or negative. The amplitude and gradient of the velocity are in time quadrature and so when velocity amplitude is most effective velocity gradient has little effect and vice versa.

## IV. APPLICATION OF THE THEORY TO THE $E$ REGION

### (a) Some Features of a "Chapman Layer"

From radio observations it is known that the  $E$  region closely resembles a "Chapman region" in which

$$I = I_0 \cdot e^{(1 - z - e^{-z} \sec \chi)} \dots\dots\dots (51)$$

where  $\chi$  is the solar zenith angle and  $z$  is the reduced height. The reduced height is measured from a datum level  $h_0$  in units of the scale height  $H$ , which is assumed to be constant in the region considered.

If the time  $t$ , measured in seconds, is reckoned from local noon

$$\cos \chi = \cos \theta \cdot \sin \delta + \sin \theta \cdot \cos \delta \cdot \cos (t/240)^\circ \quad \dots \quad (52)$$

where  $\theta$  denotes the colatitude and the seasonal parameter  $\delta$  denotes the sun's north declination.

At the equinoxes ( $\delta=0$ ) at the equator ( $\theta=90^\circ$ )

$$\chi = (t/240)^\circ \propto \omega t \quad \dots \quad (53)$$

From (5) and (51)

$$\sigma = e^{\frac{1}{2}(1 - z - e^{-z} \sec \chi)} \quad \dots \quad (54)$$

It is easily shown that  $\sigma$  has a spatial maximum

$$\sigma_m = \cos^{\frac{1}{2}} \chi \quad \dots \quad (55)$$

occurring at the height

$$z_m = \log \sec \chi \quad \dots \quad (56)$$

At noon  $\sigma_m$  has a temporal maximum. At the equator at the equinoxes, the noon value of  $\chi$  is zero giving

$$\sigma_m = 1$$

$$z_m = 0,$$

$$N_{max} \approx \eta_{max} = \eta_0.$$

Thus  $\eta_0$  is the approximate noon value of  $N_{max}$  at the equator at the equinoxes. Measurements of penetration frequency indicate that in the  $E$  region  $\eta_0$  is of order  $10^5$ .

For purposes of computation from tables of the exponential function, (54) and (55) may be expressed in the form

$$\sigma = e^{-\frac{1}{2}(e^{z_m - z} + z - 1)} \quad \dots \quad (57)$$

$$\sigma_m = e^{-\frac{1}{2}z_m} \quad \dots \quad (58)$$

In Figure 3 are  $z$  contours which show the variation of  $\sigma$  with  $\chi$  at different heights. Since  $n \approx \sigma$  these curves depict the approximate variation of the electron density with  $\chi$ . Also in Figure 3 are  $\lambda$  contours for  $k=0.3$ . A  $\lambda$  contour traces the progress of a fluid element across the  $z$  contours and also gives the variation of  $\sigma$  with both height and time.

Elimination of  $\lambda$  from

$$\sigma = \sigma(\lambda, t),$$

$$\frac{\partial \sigma}{\partial \lambda} = 0$$

gives the envelope

$$\sigma = \sigma_m = \cos^{\frac{1}{2}} \chi \quad \dots \quad (59)$$

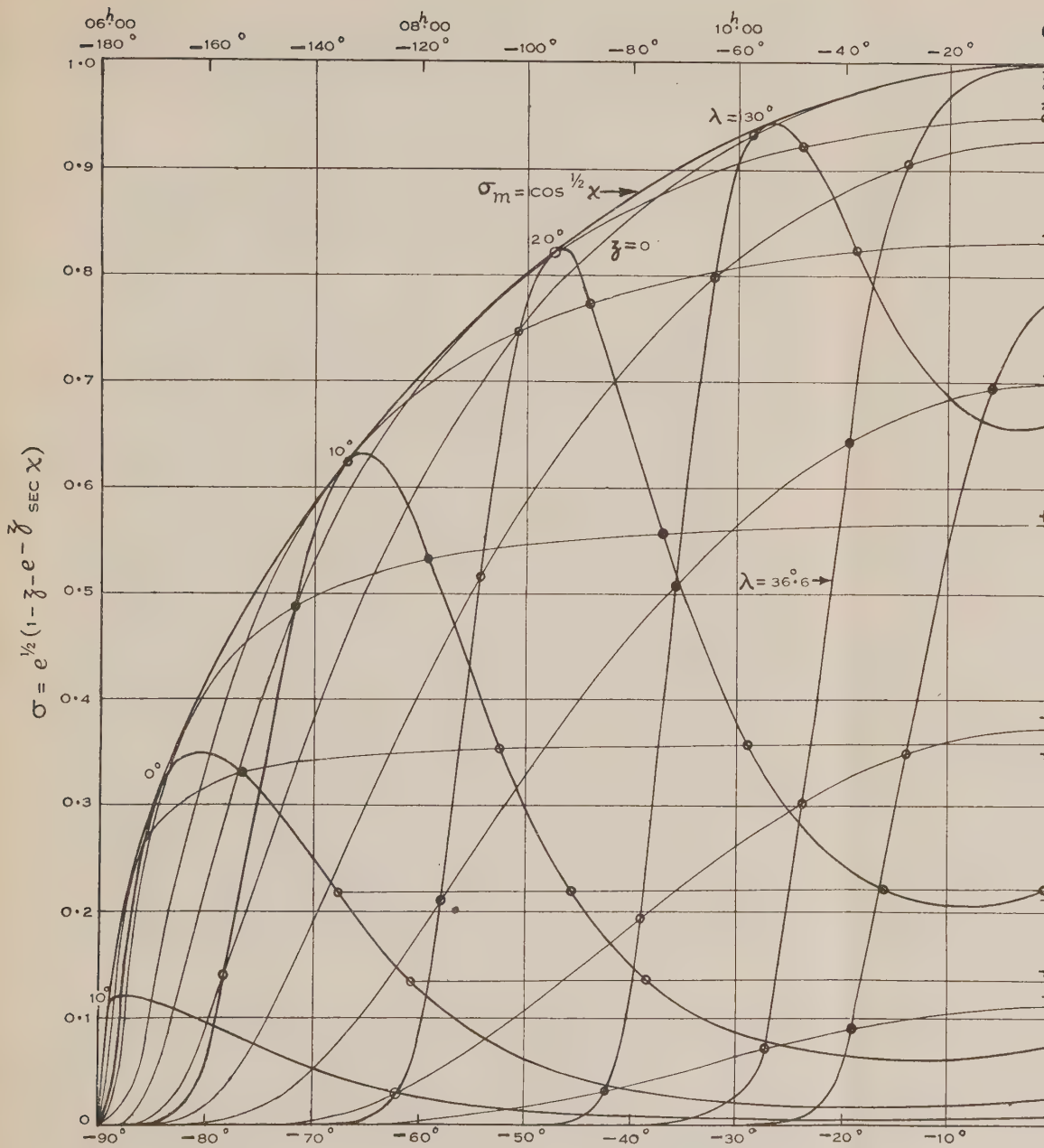
At a point of tangency on the envelope

$$z = z_m = \log \sec \chi \quad \dots \quad (60)$$

Hence a  $z$  contour touches the envelope where

$$\chi = \arccos(e^{-z}).$$

At the equator at the equinoxes  $\chi \propto \omega t$  and the curves of Figure 3 may be used for determination of " $a$ " by quadratures.


 Fig. 3.—Curves showing the diurnal variation of  $\sigma$  at the equator at the equinoxes ( $\alpha$ ) in a static



(b) *Selection of Parameters for the E Region*

In order to estimate the effect of tidal motion in the *E* region it is proposed now to investigate the diurnal variation of maximum electron density and height at the equator at the equinoxes. We shall consider a "Chapman region" which is static and the same region subject to solar tidal motion having a period of 12 hours.

Then

$$\begin{aligned}\omega &= (1/120)^\circ \text{ per sec., } (30^\circ/\text{hr.}) \\ 2\chi &= \omega t, \dots\dots\dots (61)\end{aligned}$$

Also from (53)

$$a = 2 \int_{\chi}^{\chi_1} \sigma \cdot d\chi \dots\dots\dots (62)$$

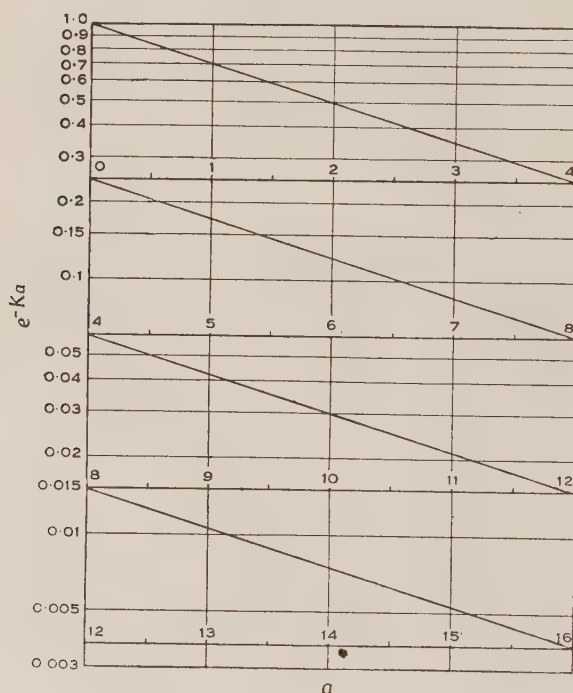


Fig. 4.—( $K=0.35$ ).

The recombination coefficient,  $\alpha$ , is of order  $10^{-8}$  and we shall take :

$$\alpha\eta_0 = 1.5 \times 10^{-3} \text{ sec.}^{-1},$$

$$K = \frac{2\alpha\eta_0}{\omega} = 0.35 \text{ per angular degree.}$$

Guided by estimates of  $v_0$  and  $\beta$  for the  $F_2$  region, we take :

$$v_0 = 2.25 \text{ per hour}$$

(i.e. 22.5 km./hr. if  $H$  is taken as 10 km. for the *E* region).

$$\beta = 4^\circ \text{ (0.07 radian).}$$

[This value of  $\beta$  corresponds to a wavelength of about 1000 km.]

$$k = \frac{\beta v_0}{\omega} = 0.3.$$

Finally we shall consider the case of upward drift during the morning and downward drift during the afternoon. These conditions apply when  $\gamma=0$ , giving

$$v = v_0 \sin(2\chi + 4z) \dots\dots\dots (63)$$

The static layer, at the equator at the equinoxes, for which  $K=0.35$  will be referred to hereafter as Region *S* and the same layer subject to ionic-drift velocity  $v$  given by (63) and  $v_0=2.25$  hr.<sup>-1</sup> will be referred to hereafter as Region *M*.

### (c) Computation of Electron Density in a "Moving Layer"

In Figure 1 is shown the variation of  $(\beta z + \gamma + \lambda)$  with  $x$  for  $k=0.1$  and  $k=0.3$ . Since an increment  $\Delta x = 2\pi$  produces a corresponding increment  $(-\delta)$  in  $(\beta z + \gamma + \lambda)$  it suffices to plot the curves for values of  $x$  in a range of  $360^\circ$ .

With the aid of the curve for  $k=0.3$  and the relations

$$2\chi = \omega t = x - \beta z - \gamma$$

corresponding values of  $z$ ,  $x$ ,  $\omega t$ , and  $\chi$  may be obtained for assigned values of  $\lambda$  and  $\gamma$  (in Region *M*,  $\gamma=0$ ).

The  $\lambda$  contours of Figure 3 are plotted by using paired values of  $z$  and  $\chi$  to locate points on the basic  $z$  contours; while a  $\lambda$  contour is being plotted  $\sigma$  may be read off and tabulated. From tabulated values of  $x$  and  $\sigma$  the corresponding values of  $\sigma'$  may be computed.

It is now necessary to select a time  $t=t_1$  at which  $n$  is to be measured, whereupon " $a$ " may be found by quadratures and tabulated as a function of  $\chi$ . Having found " $a$ " the value of  $e^{-Ka}$  may be obtained with the required accuracy from Figure 4. Measurement by quadratures gives  $n_1$  (the value of  $n$  when  $\chi=\chi_1$  and  $z=z_1$ ) as illustrated in Table 2.

A study of the curves in Figure 3 indicates that, at  $\chi=\chi_1$  the maximum value of  $n$  is associated with a  $\lambda$  contour which touches the envelope in the vicinity of  $\chi=\chi_1$ . At the point  $\chi$  on the envelope

$$x = 2\chi + \beta \log \sec \chi + \gamma$$

This value of  $x$  gives a corresponding value of  $(\beta z + \gamma + \lambda)$  from which the value of  $\lambda$  is obtained for the contour which touches the envelope at the selected point. In this way appropriate values of  $\lambda$  may be found for contours associated with values of  $n$  in the vicinity of  $n_{max}$ . The values of  $n_{max}$  and  $z_{max}$  (the height at which  $n_{max}$  occurs) are found with the aid of accurate plots of the crests of adjacent  $\lambda$  contours which touch the envelope near  $\chi_1$ . It is necessary to plot these curves between  $\chi_1$  and a  $\chi$ -value for which  $e^{-Ka}$  is very small. Figure 5 shows the  $\lambda$  contours and Figure 6 the  $(\sigma', e^{-Ka})$  curves used for determination of  $n_{max}$  and  $z_{max}$  at  $\chi=30^\circ$ . While  $n_{max}$  may be found with considerable accuracy

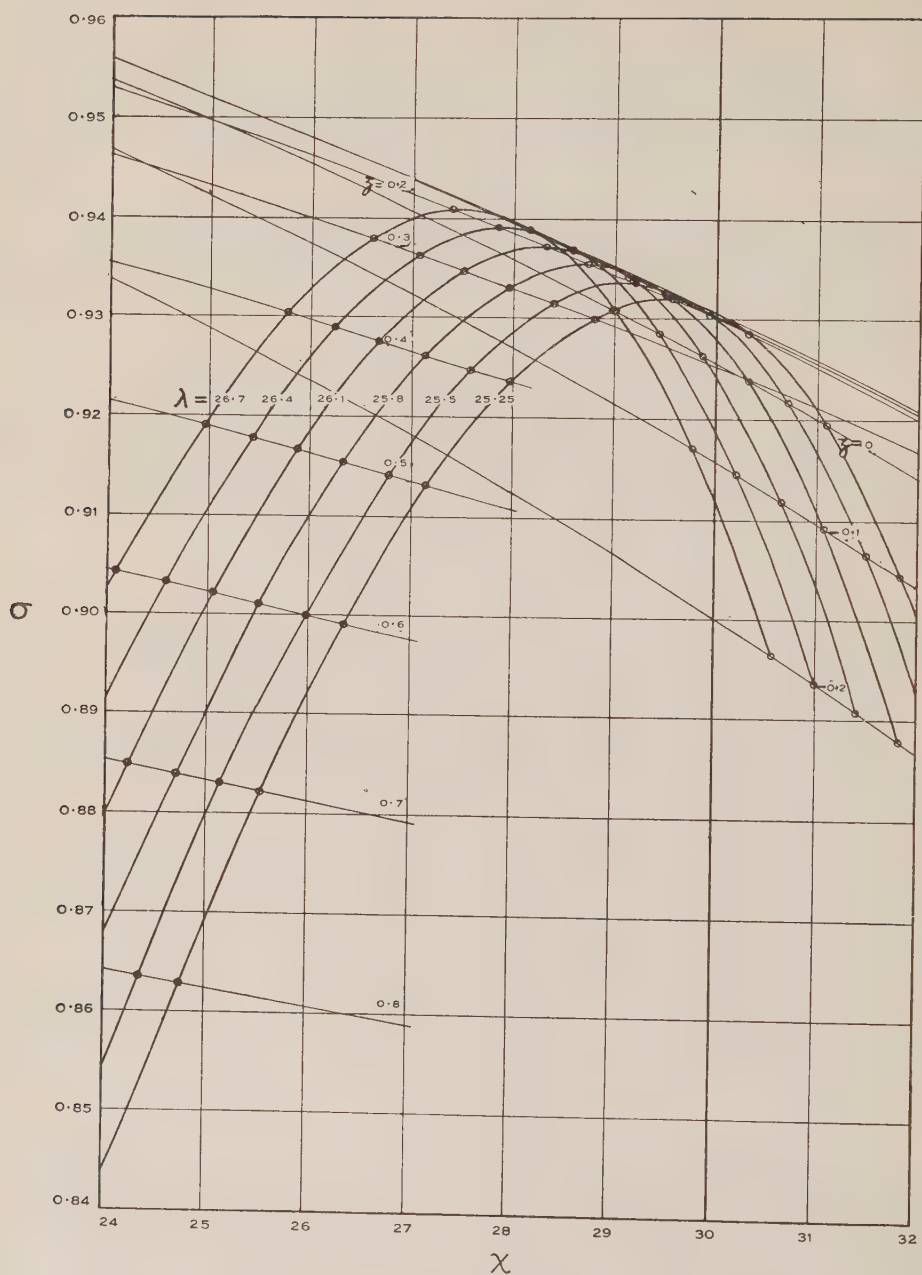


Fig. 5.— $\lambda$  contours which touch the envelope ( $\sigma_m$ ) near  $\chi = 30^\circ$ . The curve  $\lambda = 25.25$  touches at  $\chi = 30^\circ$ .

TABLE 2

$\lambda = 26^\circ.1$ $\chi_1 = 30^\circ$ $x_1 = 59^\circ.94$		$\gamma = 0^\circ$ $z_1 = -0.015$ $1 - k \sin x_1 = 0.7404$	
$z$	$\chi$	$e^{-Ka}$	$\sigma'$
-0.015	30	1	0.6844
0.1	29.11	0.56	0.6950
0.2	28.32	0.38	0.7001
0.4	26.70	0.115	0.6995
0.6	25.08	0.040	0.6871
0.8	23.36	0.014	0.6666

$$n'_1 = \int_0^1 \sigma' \cdot d[e^{-Ka}] = 0.6943$$

$$n_1 = \frac{0.6943}{0.7404} = 0.9378$$

$z_{max}$  is difficult to establish owing to the very small ionization gradient in the vicinity of  $n_{max}$ .

(d) *Computation of Electron Density in a Static Layer*

The determination of  $n$  in a static layer (see Section II) is illustrated in Table 3 for the height  $z_m$  at  $\chi = 30^\circ$  in Region  $S$ . The  $(\sigma, e^{-Ka})$  curve is shown in Figure 7.

TABLE 3

$K = 0.35$ $\chi_1 = 30^\circ$ $z = 0.144, (=z_m)$		
$\chi$	$\sigma$	$e^{-Ka}$
30°	0.9306	1
29°	0.9352	0.52
27°	0.9437	0.14
25°	0.9515	0.037
23°	0.9585	0.010

$$n_1 = \int_0^1 \sigma \cdot d[e^{-Ka}] = 0.9373$$



The maximum value of  $n$  at  $\chi = \chi_1$  is given, with sufficient accuracy, by

$$n_{max} = \int_0^1 \sigma(z_m, t) \cdot d[e^{-Ka}] \dots\dots\dots (67)$$

where

$$z_m = \log \sec \chi_1.$$

In the vicinity of a point of tangency with the envelope ( $\sigma_m$ ) a  $z$  contour of  $\sigma$  is almost coincident with the envelope curve. Consequently " $a$ " may usually be determined by measuring the area under the appropriate arc of the envelope. The height, at which  $n_{max}$  occurs ( $z_{max}$ ), is very difficult to measure exactly owing to the marked proximity of adjacent  $z$  contours as they approach

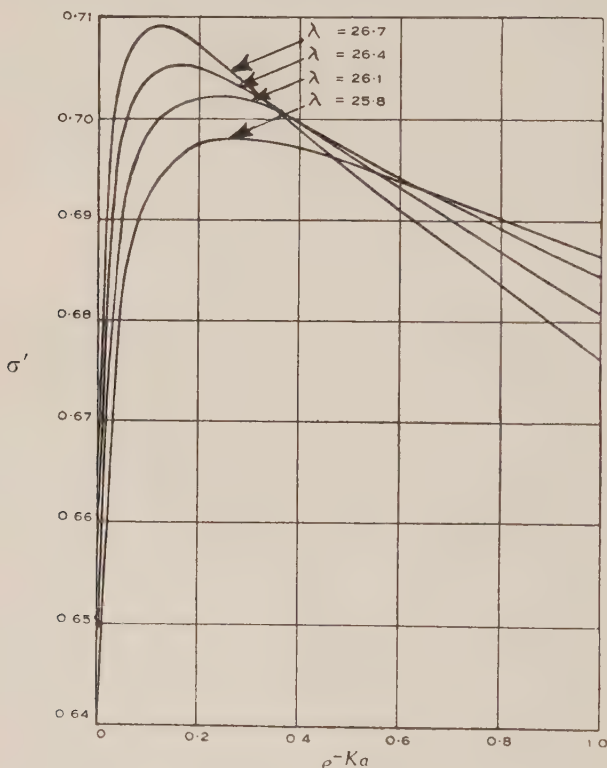


Fig. 6.—( $\sigma'$ ,  $e^{-Ka}$ ) curves for Region  $M$ ;  $\chi_1 = 30^\circ$ .

the envelope; however,  $z_{max}$  does not differ greatly from the level of maximum ion production ( $z_m$ ). The latter level has velocity

$$\frac{dz_m}{dt} = \tan \chi \cdot \frac{d\chi}{dt}$$

and is closely pursued by  $z_{max}$ , the lag being very slight until  $\tan \chi$  becomes large. Although  $z_{max} \geq z_m$  according as  $\chi \leq 0$ , it will be assumed that, during the middle part of the day,  $z_{max}$  is identical with  $z_m$  in a static layer such as Region  $S$ . In comparing values of  $z_{max}$  in a static and a "moving" layer the above assump-

tion appears justifiable if ionic drift is the dominant factor which causes  $z_{max}$  to differ from  $z_m$ .

(e) *Comparison of the "Moving" Layer (M) and the Static Layer (S)*

The graphical and numerical methods outlined above have been used to determine  $n_{max}$  and  $z_{max}$  for the "moving" layer  $M$  and the static layer  $S$  between 8 a.m. and 4 p.m. In Table 4 are shown values of  $n_{max}$  and  $z_{max}$ . Also tabulated is  $\sigma_m$  (the first approximation to  $n_{max}$ ).

$N_{max}$  and  $h_{max}$  are readily found from the relations

$$N_{max} = \gamma_0 \cdot n_{max}$$

$$h_{max} = h_0 + H \cdot z_{max}$$

TABLE 4

Time	$\chi$	$\sigma_m$	Region $M$		Region $S$	
			$n_{max}$	$z_{max}$	$n_{max}$	$z_m$
8.00 a.m.	-60	0.7071	0.656	1.05	0.682	0.693
9.20 a.m.	-40	0.8753	0.852	0.52	0.864	0.267
10.40 a.m.	-20	0.9694	0.971	0.215	0.964	0.062
11.20 a.m.	-10	0.9924	1.001	0.09	0.990	0.015
12 noon	0	1	1.015	0	0.9995	0
12.40 p.m.	10	0.9924	1.008	-0.06	0.994	0.015
1.20 p.m.	20	0.9694	0.983	-0.04	0.974	0.062
2.00 p.m.	30	0.9306	0.938	-0.015	0.937	0.144
2.40 p.m.	40	0.8753	0.878	0.05	0.885	0.267
4.00 p.m.	60	0.7071	0.707	0.45	0.726	0.693

The significance of the results is made evident by the curves of Figure 8, which portray the departure from normal "Chapman" behaviour occasioned by the presence of ionic drift. Drift velocity is represented by the variation of  $v$  at the datum level  $h_0(z=0)$ . In Region  $M$ ,  $N_{max}$  and  $h_{max}$  exhibit a definite departure from their "Chapman" values in Region  $S$ . The curves show the percentage departure of  $N_{max}$  and the departure in km. of  $h_{max}$  when  $H$  is taken equal to 10 km.

It is seen that  $\Delta h_{max}$  (the departure of  $h_{max}$ ) may be as much as 2 to 3 km. and closely follows the variation in drift velocity; thus, in accord with simple physical considerations,  $h_{max}$  is elevated when there is upward drift of ions and depressed when there is downward drift.

Departure in  $N_{max}$  is as high as 2 per cent. to 4 per cent. and clearly its variation is not simply related to the drift velocity. Nevertheless, it is not difficult to interpret the departure curve for  $N_{max}$ . General considerations indicate that velocity amplitude causes a decrease in  $N_{max}$ . In the late morning

and early afternoon the drift velocity is small in Region  $M$ , and very little decrease occurs, but a large positive velocity gradient produces condensation of ionization and a consequent increase in  $N_{max}$ . Clearly velocity gradient plays a major part in shaping the departure curve for  $N_{max}$ .

(f) *Discussion of the Effect of Drift Velocity Parameters*

It has been pointed out that the velocity amplitude effect causes a decrease in  $N_{max}$ . The general effect of velocity gradient may be roughly estimated by manipulation of the factor  $(1 - k \sin x)/(1 - k \sin x_1)$  which it introduces into the integrand of the expression for  $n_1$ , viz.

$$n_1 = \int_0^1 \frac{1 - k \sin x}{1 - k \sin x_1} \cdot \sigma(\lambda, t) \cdot d[e^{-Ka}].$$

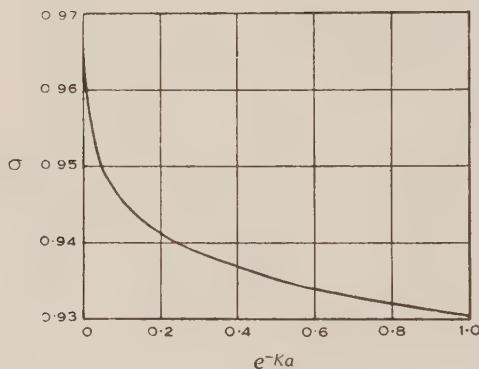


Fig. 7.—( $\sigma, e^{-Ka}$ ) curve for Region  $S$ ;  $\chi = 30^\circ$  and  $z = z_m = 0.144$ .

The quantity  $(x_1 - x)$  is positive and is usually small giving

$$\frac{(1 - k \sin x)}{(1 - k \sin x_1)} \approx 1 + k \cos x_1 \cdot \frac{\sin(x_1 - x)}{(1 - k \sin x_1)}.$$

Hence velocity gradient introduces a term proportional to and having the same algebraic sign as  $k \cos x_1$ . Consequently a change of phase ( $\gamma = 180^\circ$ ) or a reversal of the phase velocity ( $\omega/\beta$ ) by making  $\beta$  negative would give rise to a decrease in  $N_{max}$  instead of the increase exemplified by the central portion of the  $N_{max}$  departure curve in Figure 8. Further, a reversal of both phase and phase velocity would not affect  $k \cos x_1$  and would therefore produce no marked change in  $N_{max}$ . On the other hand  $\Delta h_{max}$  appears to follow closely the variation of drift velocity (as in Fig. 8). Rough quantitative tests indicate that if  $\beta$  changes sign the magnitude of  $\Delta h_{max}$  is affected but no reversal of sign occurs. Reversal of phase, however, causes approximate reversal of  $\Delta h_{max}$ ; consequently an experimental departure curve for  $h_{max}$  should indicate approximately the phase ( $\gamma$ ) of the drift velocity.

(g) *Discussion of the Results*

The results show that vertical drift of ions causes a small perturbation of the  $E$  region. It seems reasonable to suppose that the main features of the results obtained for a region at the equator at the equinoxes are applicable to other latitudes and seasons provided the global distribution of the phase constant ( $\gamma$ ) is taken into account. The method of solution precludes accurate deter-

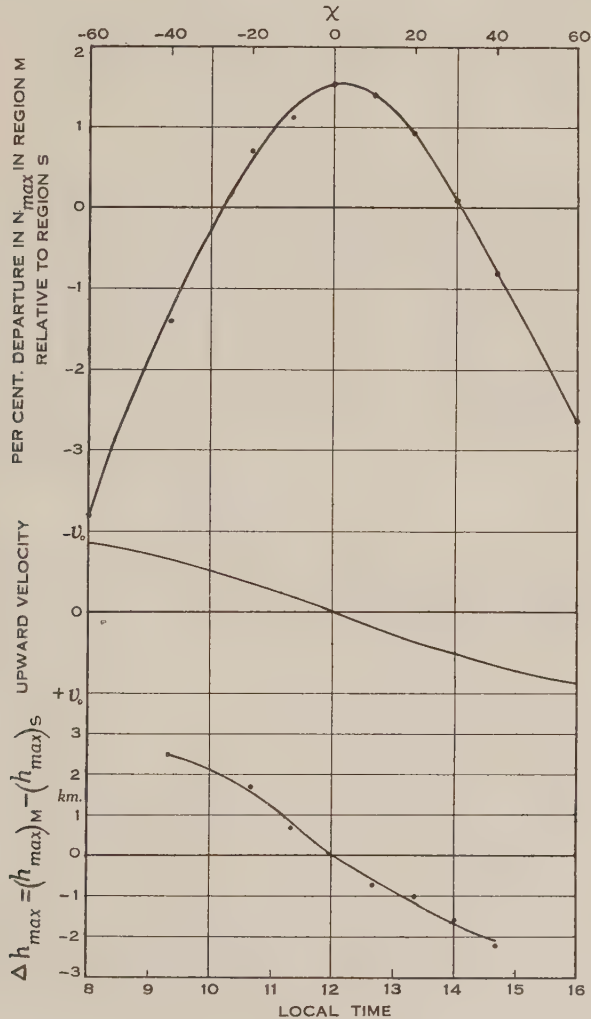


Fig. 8.—Curves illustrating the departure of Region  $M$  from normal "Chapman" behaviour.

mination of electron density near sunrise and sunset. It should be possible, however, to investigate a "moving" layer near sunrise and sunset by use of Taylor's theorem after the manner described by Millington. Such a procedure would be laborious and hardly worth while unless the results were required for a



specific purpose. From a study of the night-time solution (44) and Figure 2 it is clear that, if recombination is appreciable, ionic drift exerts only a small influence in a region at night-time. The effect of ionic drift when recombination is negligible ( $F_2$  region) has been discussed by Martyn.

Analysis of ionospheric data for a large number of years has revealed in the  $E$  region departures from "Chapman" values, which are readily explained by the above theoretical results. Comparison of theoretical and experimental data has yielded information about the phase of the atmospheric solar tides, knowledge of which is useful for locating the currents causing the solar magnetic variations. A full discussion of these applications of the results is given elsewhere by Martyn.

#### V. ACKNOWLEDGMENTS

This investigation was suggested to the author by Dr. D. F. Martyn, whose keen interest and helpful advice greatly stimulated the progress of the work. The author is indebted to the Radio Research Board, C.S.I.R., and to the Ionospheric Prediction Service of the Commonwealth Observatory for providing facilities. Particular thanks are due to Misses O'Dwyer and Farrell of the Radio Research Board staff for assistance with the computations. It is a pleasure to acknowledge the constant encouragement of Sir John Madsen. The work is published by permission of the Director of Technical Education.

#### VI. REFERENCES

- (1) CHAPMAN, S.—*Proc. Phys. Soc. Lond.* **43** : 26 (1931).
- (2) MARTYN, D. F., and PULLEY, O. O.—*Proc. Roy. Soc. A* **154** : 455 (1936).
- (3) MILLINGTON, G.—*Proc. Phys. Soc. Lond.* **44** : 580 (1932).
- (4) MILLINGTON, G.—*Ibid.* **47** : 263 (1935).
- (5) MARTYN, D. F.—*Proc. Roy. Soc. A* **189** : 241 (1947). *Ibid.* **190** : 273 (1947).

# RADIO SUPERREFRACTION IN THE COASTAL REGIONS OF AUSTRALIA

By F. J. KERR\*

[*Manuscript received September 27, 1948*]

## *Summary*

A survey has been made of the incidence of radio superrefraction round the coast of Australia and nearby islands. It is based on the systematic reporting of the strengths of echoes from ground objects observed by Royal Australian Air Force 200 Mc/s. radar stations. Observations taken hourly over periods of up to eighteen months have been analysed in terms of geographical position, season of year, time of day, and synoptic weather situations. The main points arising from the analysis are : (i) The frequency of occurrence of superrefraction at different times and places has been determined ; (ii) superrefraction is strongest and most regular in north-west Australia (occasional very long-range echoes occur in this region) ; (iii) superrefraction in coastal Australia is associated with one general condition, the offshore movement of a relatively warm and dry air mass ; and (iv) superrefraction on inland echoes is due almost exclusively to night-time radiation inversions.

## I. INTRODUCTION

When ultra-high-frequency radio waves were first used, it was expected that they would resemble light waves in propagation characteristics. Before the war it had become apparent that under some conditions quite high field strengths could be received well beyond the horizon.

The concentration of effort on the ultra-high frequencies brought about by the introduction of radar soon showed that increased field strengths beyond the horizon were quite common. Owing to its tactical importance, the subject of "anomalous propagation", as it was called, then received intensive study. The main types of meteorological conditions under which the increased bending of waves ("superrefraction") was likely to occur were identified, and the subject was put on a sound basis by Booker, who developed a theory applicable to the condition of greatest importance in England.

The continuous operation of widely-dispersed radar stations during the war provided an opportunity for a study of ultra-high-frequency propagation conditions over a wide area, and a large-scale study of superrefraction was begun early in 1944 by the Council for Scientific and Industrial Research working in collaboration with the Royal Australian Air Force.

In this investigation a large field of space and time was covered at the expense of accurate and detailed observations. From the nature of the reporting system the data were semi-quantitative and only broad statistical conclusions

\* Division of Radiophysics, C.S.I.R.

were drawn. Nevertheless, it is possible to establish the general trends in the variation of the incidence of superrefraction and its relation to meteorological conditions.

It was demonstrated, for example, that superrefraction in summer often extends over a large area of southern Australia simultaneously. The effects of superrefraction are most striking in north-west Australia, where echoes from Timor (400 miles away) are received fairly frequently in some seasons and echoes at somewhat smaller ranges recur with great regularity at almost the same time day after day.

The first general review of the radio-climatology of Australia was given by Booker(1) after a short visit to this country. The present account begins with a discussion of the evidence upon which the survey is based. Then follow the main results of the analysis, presenting a picture of the radio-meteorology of the region. Full records of the observations and detailed reports on various aspects of the analysis are available in this Laboratory.

## II. EVIDENCE UPON WHICH THE SURVEY IS BASED

Reports were received from 112 Royal Australian Air Force 200 Mc/s. air warning radar stations located at points near the coast of Australia and neighbouring islands. For the purpose of analysis, the stations were divided on climatological grounds into eleven regional groups. Attention was concentrated on a selected list of about twenty-five stations which had suitable echoes available, made reliable observations, and gave a wide geographical distribution. The positions of these stations are marked on the map in Figure 4. The period covered was March 1944 to August 1945, though some stations closed down or operated on restricted hours towards the end of this period. Detailed meteorological information was supplied by the Directorate of Meteorological Services, Royal Australian Air Force.

Each radar station selected up to eight echoes from fixed topographical features, below the radio horizon during periods of normal propagation. The strengths of these echoes were observed every hour. The distances involved were in most cases between 50 and 150 miles, but some stations in the north also kept records of echoes from distances of 300-400 miles.

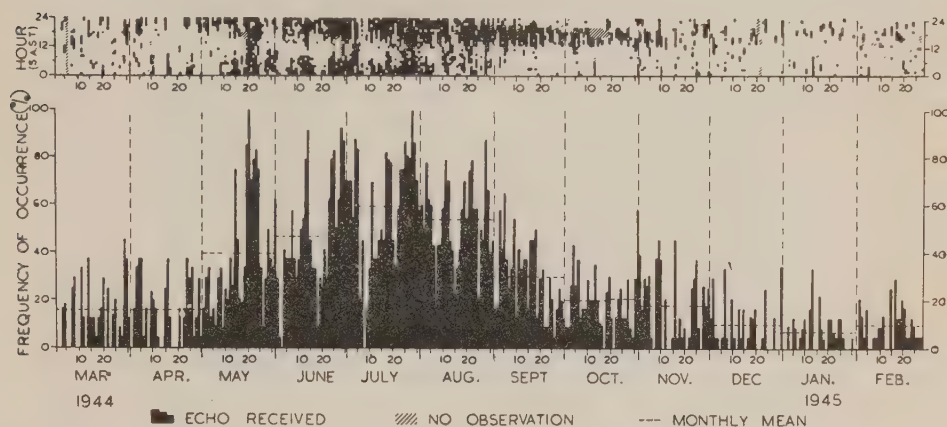
The assessment of propagation variations was in general made from the appearance or non-appearance of the echoes. The observations of actual echo strength covered only a very limited range (ten times in voltage from threshold to saturation), which was small in comparison with the changes in field strength known to take place. Superrefraction has thus been treated as essentially an "on-off" effect.

The frequency of occurrence of superrefraction has been assessed from the percentage of observations in which a given echo was seen. This percentage gives a "frequency of occurrence" of the echo under the conditions considered, whose daily (day-to-day) and diurnal (hour-to-hour) variations have been calculated for each echo for each month. The mean frequency of occurrence for each of successive months has been used in the examination of seasonal variations.

These quantities have been worked out in many cases for all the echoes at a station lumped together, and for groupings of sea-path, land-path, and shore-path echoes. Having established the normal behaviour for each echo in respect to superrefraction, it is possible to correlate particular deviations from normality with particular synoptic conditions.

**STATION 59, LEE PT., DARWIN, HT. 230 FT.**

ECHO 3—322°, 44 MILES (SEA-PATH).



**STATION 33, CAPE NATURALISTE, HT. 680 FT.**

ECHO 2—015°, 110 MILES (SEA-PATH).

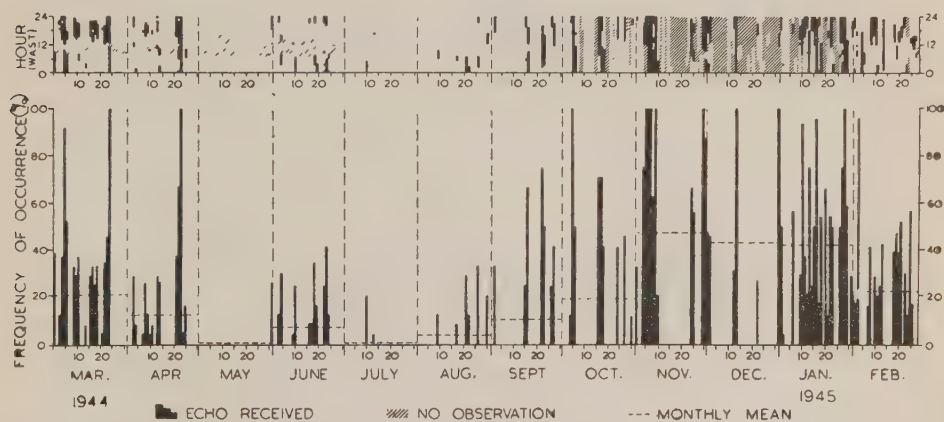


Fig. 1.—Typical occurrence diagrams for one year (March 1944–February 1945) for two echoes. The upper diagram of each pair presents the hourly observations throughout the period, and the lower diagram the “frequency of occurrence” for each day.

The detailed results of one year’s observations on two typical echoes are given in Figure 1,\* where the upper diagram of each pair presents the hourly observations through the period, and the lower diagram the “frequency of occurrence” for each day.

\* E.A.S.T., S.A.S.T., W.A.S.T., the abbreviations used for standard times in the figures, refer to Eastern Australian, South Australian, and Western Australian Standard Time.



The survey has several limitations arising from the nature of the observational data. First, no direct comparison can be made between propagation conditions at different stations. An attempt was made to get an absolute measure of echo field strengths from estimates of the echoing powers of the reflecting objects, but the uncertainties were too great and little success was achieved. Again, echoes from well out to sea were obtainable in a few localities only, most of the echoes having trajectories over land or near the shore. Some information on echoes from ships was available but this was sporadic.

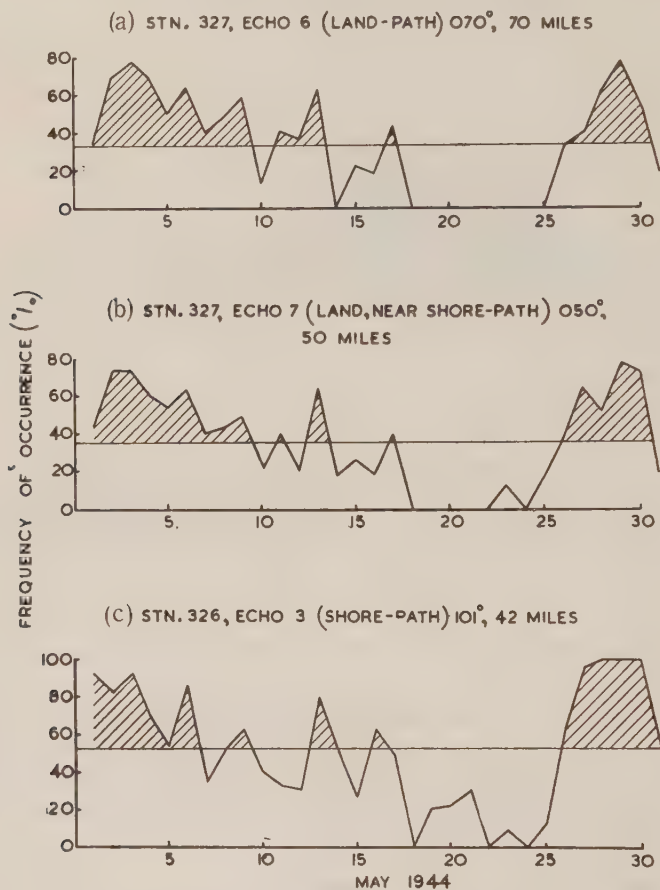


Fig. 2.—Typical results from nearby echoes and stations, illustrating the internal consistency of the data. (Station 326 is  $022^{\circ}$ , 120 miles from station 327.) The horizontal line through each curve indicates the monthly mean.

The observations were made by Service operators, with frequent changes of personnel and often under difficult conditions. Small variations in equipment performance were to some extent recorded but could not be satisfactorily allowed for in working out echo occurrences. The internal consistency of the reduced data gives a measure of the reliability of the observations. Results for different



echoes at one station show a considerable measure of agreement when the echoes have similar trajectories, a sample comparison being given in Figures 2 (a) and 2 (b). Likewise the curves for neighbouring stations agree reasonably well as illustrated by Figures 2 (b) and 2 (c). Comparisons of this type, together with the fact that in many cases the variations for a particular echo show marked similarities from day to day, suggest that the main features of the curves represent physical realities, differences between observers and variations in equipment having only minor effects. Agreement is by no means complete, however, in respect to fine detail, and many of the differences may be due to subjective factors. As is common in many meteorological problems, it has been necessary to consider a significantly large number of cases to establish each conclusion.

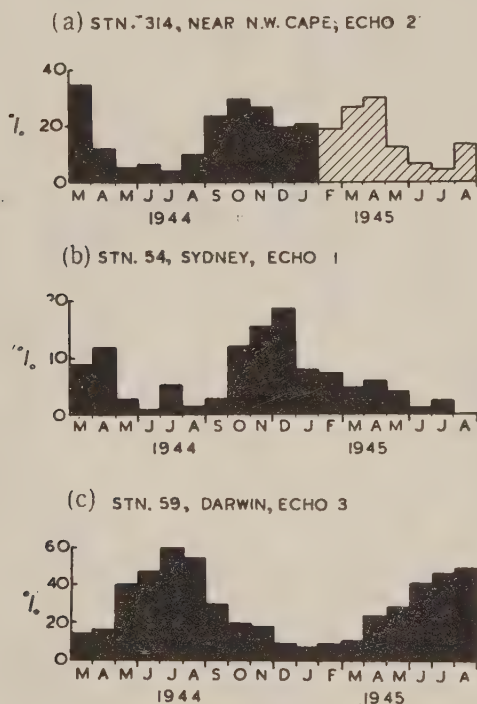


Fig. 3.—Typical seasonal variations (unsmoothed) covering 18 months. Hatching indicates limited hours of observation.

### III. INCIDENCE OF SUPERREFRACTION IN AUSTRALIA

(i) *Seasonal and Geographical Distribution.*—The seasonal variation of occurrence of each echo has been obtained from a plot of the monthly means; three examples, taken from different regions, are shown in Figure 3. The results for individual echoes at all the "first choice" stations are presented on a smaller scale in Figure 4, ranged round a map of Australia and New Guinea. In this case the curves have been smoothed by taking overlapping three-monthly

averages, and for convenience have been restricted to a single year. The seasonal results for sea and shore-path echoes have been summarized in idealized form in Figure 5, where the periods of maximum and minimum occurrence are indicated on "clock" diagrams, each referring to the adjacent portions of the coast.

A study of Figure 5 shows that the seasonal behaviour on the north coast near Darwin is almost exactly opposite to that in the south, the former showing a maximum in winter, the latter a maximum in summer. The transition between

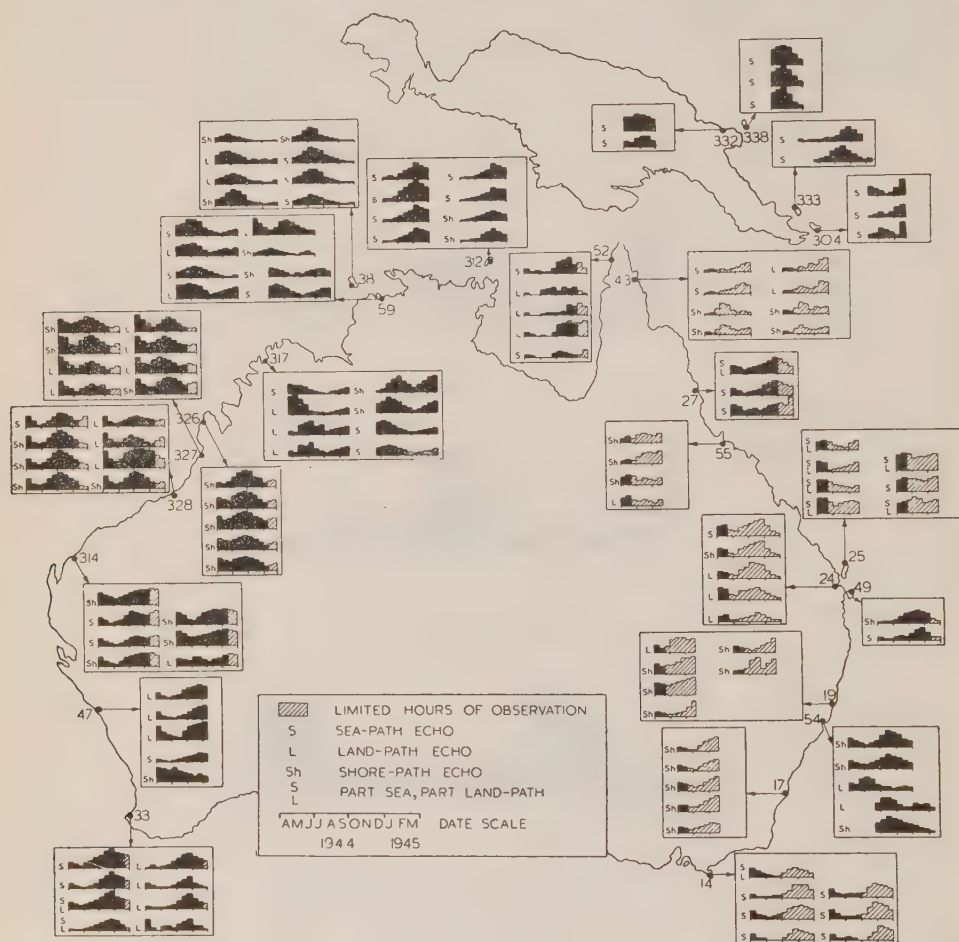


Fig. 4.—Seasonal variations for echoes at first choice stations, covering one year of operation. The data have been smoothed by taking three-monthly running averages.

the two takes place differently west or east round the coast from Darwin. To the west, in the region of most intense superrefraction, the single winter maximum found at Darwin is replaced by maxima in autumn and spring. Southwards down the west coast, the times at which these two maxima occur move towards the summer period, leading to the single summer maximum on the south coast.

Eastwards from Darwin, through the region of weak superrefraction, the time of the maximum simply changes through spring to summer.

As mentioned earlier, it has not been possible to obtain a quantitative comparison of the strength of superrefraction occurring in the different regions. However, a qualitative comparison has been made and the results are included in Figure 5, where circles of three sizes indicate the strength of superrefraction. The most intense superrefraction occurs along the north-west coast. It is fairly intense at times in any portion of southern Australia, but is never strong in

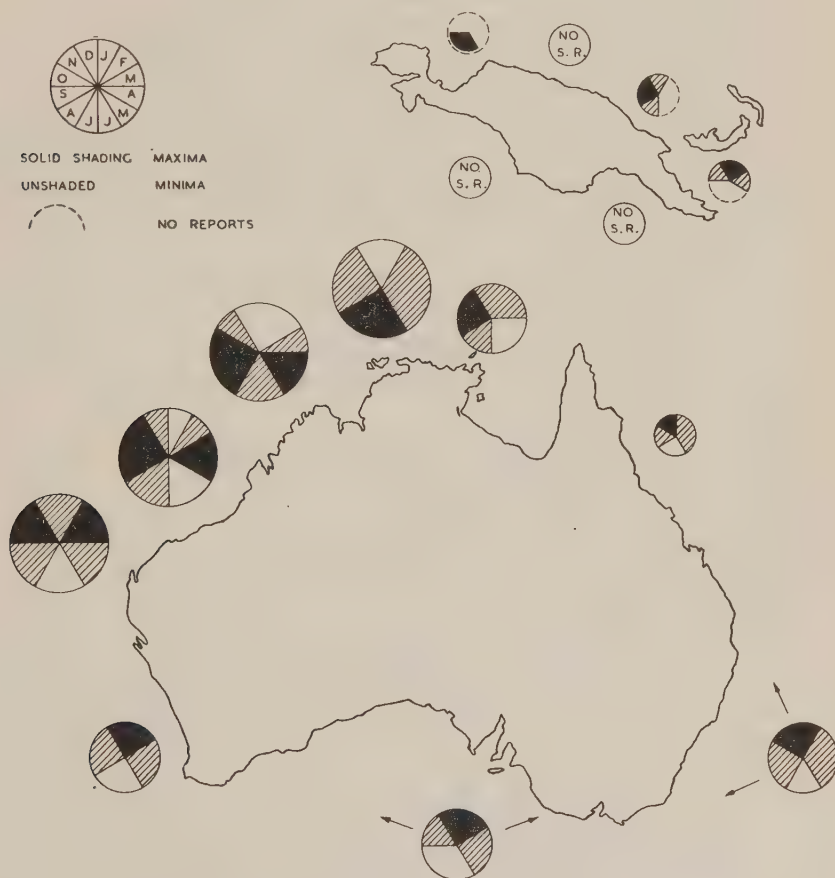


Fig. 5.—Idealized seasonal variation of coastal superrefraction, with the strength of superrefraction in each region indicated by the size of the circle of the "clock" diagram. (New Guinea has been displaced northward for convenience.)

north-east Australia, where infrequent weak superrefraction occurs. Moderate superrefraction occurs in eastern New Guinea and presumably in Geelvink Bay, but elsewhere in New Guinea little or no superrefraction occurs at the radio frequency under consideration.

(ii) *Diurnal Variation*.—During the analysis of the observations, diurnal variations were plotted for each echo for each month. The results were too

diverse, however, to be summarized in a simple form as with the seasonal variation. Examples will be given of the main types of diurnal variation for each type of echo trajectory, a more complete discussion following later in terms of the associated meteorological situations.

Diurnal variations are found to be most pronounced in the north and north-west, moderate on the east coast, and least pronounced in the south-west and south. In the north and north-west, diurnal effects recur with great regularity from day to day; this is to be expected in view of the pronounced diurnal effects which are a feature of tropical meteorology.

Sea and shore-path echoes are found to have their diurnal maximum in the afternoon, evening, or early morning, according to the region and the season. Examples are shown in Figures 6 and 7. The time at which superrefraction commences is found in general to be earlier in northern than in southern Australia, and earlier in summer than in winter, i.e. it is earlier when the air-sea temperature excess is greater. Three main types of diurnal variation can be distinguished, average figures for the periods in which superrefraction occurs being 1300–1900, 1700–0100, 2400–0900 hours.

Land-path echoes display a well-marked nocturnal maximum and day-time minimum. The time at which superrefraction begins in the early evening appears to be substantially independent of region and season. In some cases a double maximum is found, with a minimum soon after midnight. Sample curves are shown in Figure 8.

(iii) *Relation to Synoptic Systems.*—Superrefraction is often found to occur over large areas simultaneously. Booker(1) first described how, as an anticyclone moves eastward, a sector of superrefraction will detach itself from the semi-permanent sector on the north-west coast and move across from the west coast round the south and east coasts ahead of the succeeding polar front, being finally obliterated on the east coast as the front advances further.

The relation between superrefraction and synoptic systems has been studied from mass-plots of echo occurrence. Examples for three stations in southern Australia are shown in Figure 9, where the lumped echo occurrence is plotted against position in an idealized anticyclone. Each case covers twenty days, during which three or four anticyclones passed eastward over the station. From such plots it appears that superrefraction in southern Australia occurs most frequently in the rear halves of the migratory anticyclones, ceasing at a given place as the polar front, which precedes the next anticyclonic cell, passes. A typical situation is shown in Figure 10.\* The eastward movement of the sector of superrefraction is most evident in summer and least in winter. (The polar fronts and anticyclonic cells themselves are most clearly defined in summer.) The variation is, of course, more complicated when the synoptic system is more complex than a simple succession of well-defined anticyclones. Owing to coastal front phenomena, the sequence of events is rather more complex in the south-west corner (e.g. at station 33) than elsewhere in the south. The characteristic feature

\* The air-mass symbols are: T<sub>c</sub>, tropical continental; S<sub>m</sub>, southern maritime; T<sub>p</sub>, tropical Pacific.



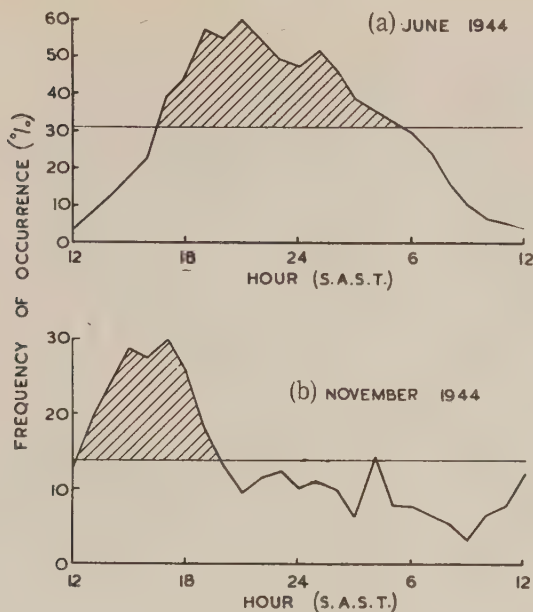


Fig. 6.—Diurnal variations, station 59 (Darwin), lumped sea-path echoes.

- (a) Superrefraction associated with offshore streaming.  
 (b) Superrefraction associated with sea breeze.

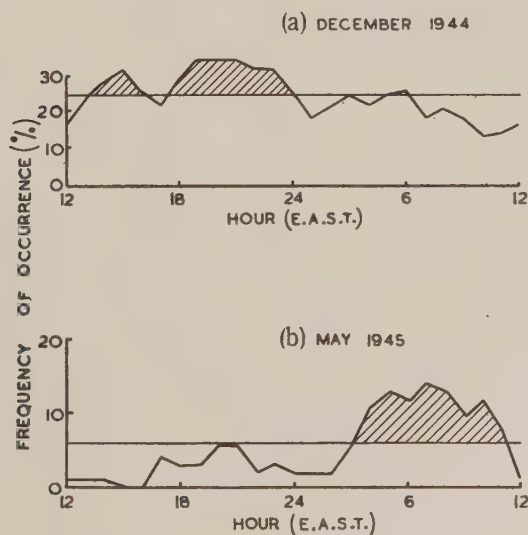


Fig. 7.—Diurnal variations, station 54 (Sydney), lumped shore-path echoes.

- (a) Superrefraction associated with coastal front and offshore streaming.  
 (b) Superrefraction mainly associated with nocturnally-cooled air carried over sea.



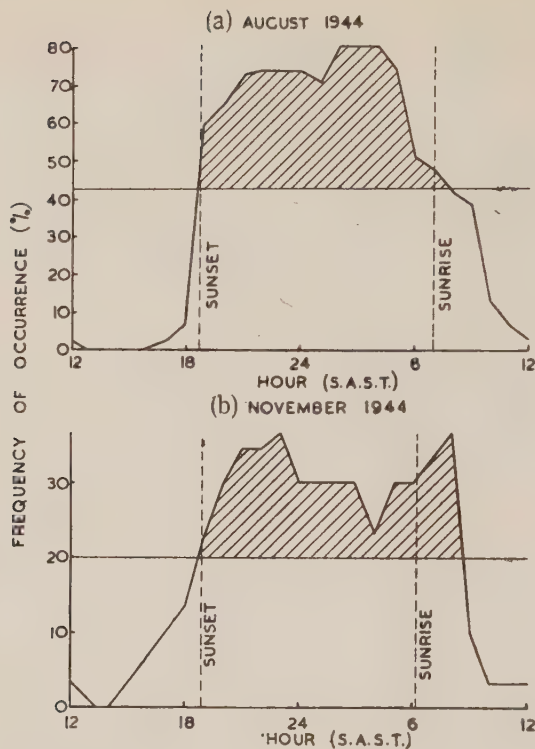


Fig. 8.—Diurnal variations, station 59 (Darwin), echo 4 (land-path).  
(a) Dry season. (b) Wet season.

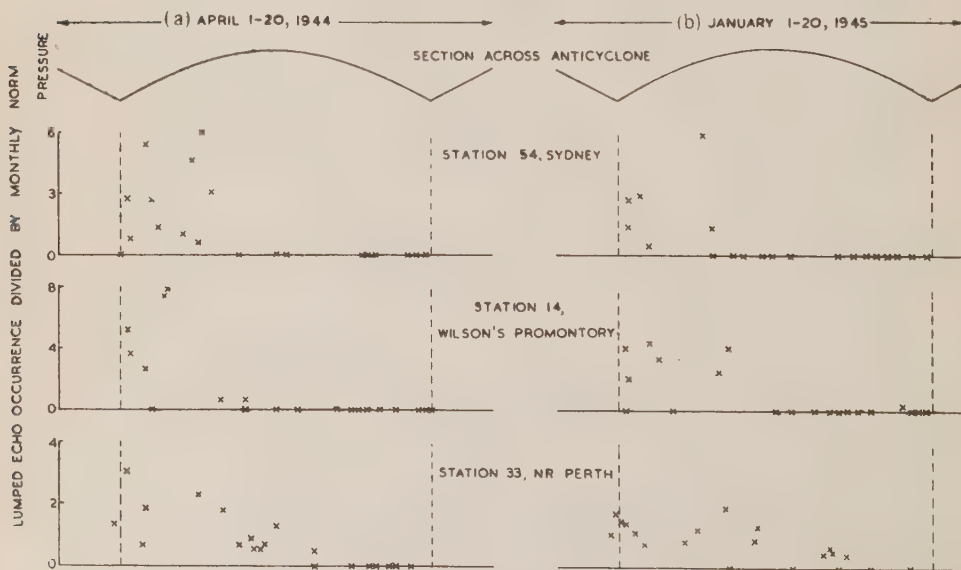


Fig. 9.—Relation between superrefraction and synoptic systems at three stations in southern Australia, for two periods of 20 days each.

of the region in the rear of an anticyclone in southern Australia is an offshore gradient wind, which carries warm, dry continental air across the coastline. Thus superrefraction in southern Australia is associated with an offshore gradient wind. In northern Australia no clear-cut relation with synoptic systems has been found, diurnal effects being more important and local variations quite large. With the very long-range echoes described in Section V, superrefraction phenomena have been found to occur over a large area simultaneously, but this is generally not the case with the echoes of smaller ranges.

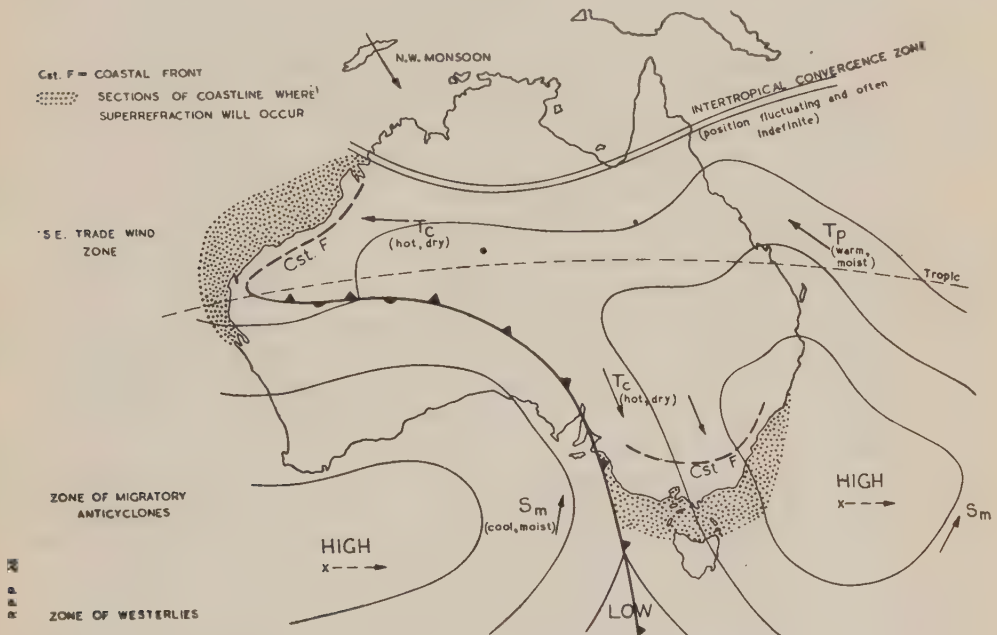


Fig. 10.—Typical meteorological situation in Australia in summer (idealized). In winter the systems move northward, the trade-wind zone extending over the whole northern coast.

#### IV. PROCESSES LEADING TO SUPERREFRACTION

The normal downward bending of radio waves in the lower atmosphere increases and tends to enhance propagation to great distances whenever the lapse of temperature with height is below normal or reversed in sign (temperature inversion) and/or the decrease of humidity with height is greater than normal. Humidity effects are generally considered to be the more important in producing superrefraction.

The classical theory for superrefraction over the sea was developed by Booker(2). It is particularly applicable to English conditions where superrefraction is generally associated with offshore advection of warm, dry air. Martyn and Squires(3), analysing early Royal Australian Air Force data, showed that coastal front phenomena are the principal cause of superrefraction in the warmer continent of Australia. Thus most superrefraction in Australia is associated with an onshore wind at the surface rather than with an offshore wind. The upper air movement, however, is offshore. Superrefraction appears on

sea paths when warm, dry air from the land overlies relatively cool and moist air. The latter may be provided in three principal ways, which depend on temperature and wind conditions and appear to be associated with the three distinctive forms of diurnal variation.

(i) *Sea Breeze, Coastal Front*.—The first type of diurnal variation is found to be associated with a large air-sea temperature excess, and an onshore surface wind at the coast.

When the temperature of the land exceeds that of the sea by more than a few degrees the temperature difference is sufficient to produce a sea breeze which can reverse the direction of a previously-existing offshore surface wind. The temperature excess at which the transition to sea-breeze conditions occurs depends on the strength and direction of the general wind, and under favourable conditions in Australia a sea breeze continues throughout the day and night, tending to form a persistent air-mass boundary near the coast.

Australian meteorologists have developed the concept of the "coastal front" to describe a phenomenon of special importance in Australia(4). A front is likely to be set up when a stream of hot tropical continental air is crossing the coast (and hence the isotherms) at a fairly small angle. That portion of the air mass which is over the land becomes warmer than that over the sea, and sea-breeze processes assist in the formation of a frontal discontinuity along or near the coastline, with a wedge of cool maritime air undercutting the continental stream. The coastal front shows a diurnal oscillation, moving inland during the day and retreating to the coast at night.

The best example is found along the west coast of Australia for a large part of the summer. A somewhat similar phenomenon occurs along the north-west, south, and east coasts in summer, when wind conditions are suitable. The sea-breeze circulation is very well developed on these parts of the Australian coast, but there is a difference of opinion amongst meteorologists as to whether the surface of discontinuity occurring in this circulation can be described as a front. Such an extension of the generally-accepted terminology seems justified for our purpose, however, as the phenomenon is essentially similar to that found in a true front in relation to superrefraction, which emphasizes temperature and humidity conditions rather than dynamical considerations.

In a typical southern Australian coastal front, the surface of discontinuity approaches ground level 50–100 miles inland, rising towards the sea. The height at the coastline is commonly 2000–3000 feet. The conditions over the sea are not well known, but a small number of aircraft flights from Sydney have shown that the seaward extent of the structure is great. Detailed investigations of local sea breezes in other regions have been carried out by Hatcher and Sawyer(5) at Madras and by Craig, Katz, and Harney(6) on the American east coast.

Offshore streaming of hot, dry air occurs typically for two to four successive days in summer in southern Australia. On the first day the sea-breeze effect is small, lasting for a short time only, but it increases in extent and duration on the succeeding days. By the second or third day the coastal front is able to maintain its existence overnight. It then lasts until the conditions which set

it up are destroyed by the arrival of a polar front. During these days the coastal front tends to move inland by day and towards the coast by night. Its intensity is greatest by day.

Superrefraction associated with a small-scale sea breeze is found to commence soon after the breeze sets in, ceasing when it dies away (see Fig. 6 (b)). With a coastal front in southern Australia, the diurnal variation of superrefraction is less (see Fig. 7 (a)), since once a coastal front is established superrefraction continues almost uniformly for days at a time. In the north-west the coastal front is a semi-permanent feature, but here the diurnal variation is greater owing to the tropical location.

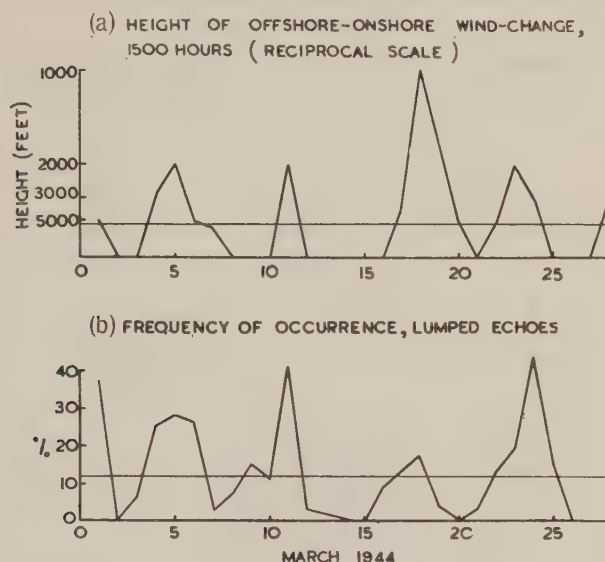


Fig. 11.—Relation between occurrence of superrefraction and height of coastal front, station 47 (Geraldton), March 1944, lumped echoes. (From Fig. 3, Martyn and Squires(3).)

On the west coast of Australia, the coastal front is the dominant cause of superrefraction in summer. A very close relation exists between abnormal echo occurrence and the height of the coastal front, as indicated by the height at the coast of the wind-change from an offshore to an onshore direction. This is illustrated in Figure 11. Insufficient data are available to determine whether the height itself is the significant quantity, or whether an increase in the intensity of the front is associated with a decrease of height.

(ii) *Offshore Streaming, Diffusion Duct*.—As warm, dry air passes over a slightly cooler sea surface, the lower layers are progressively modified by downward eddy conduction of heat to the cooler sea and upward eddy diffusion of moisture from the sea into the air. Modification of the air in the lowest layers increases the lapse-rate of refractive index, producing a surface duct for some distance out to sea. The stability of the lower air prevents destruction of the duct by convection. The resulting superrefraction commences a few hours



after the time of maximum heating some distance inland, and ceases some hours after sunset. This corresponds to the second type of diurnal variation noted in Section III (ii) above and illustrated in Figure 6 (a). Superrefraction is found to commence earlier and last longer as the temperature excess increases. This type of superrefraction occurs in northern Australia in winter and in southern Australia in spring and autumn.

(iii) *Nocturnally-cooled Air Carried Out to Sea.*—In the winter months coastal superrefraction in southern Australia is weak and infrequent, but by no means negligible. Most of the superrefraction occurs in the early morning hours and appears to be associated with radiation inversions over the land near the coast.

The diurnal variation of this type of superrefraction is shown in Figure 7 (b) for echoes whose paths lie along the shore up to seven miles out from the coastline. The maximum occurrence is later in the night than that due to simple advection (Fig. 6 (a)) or to radiation inversions over land (Fig. 8). This mechanism is

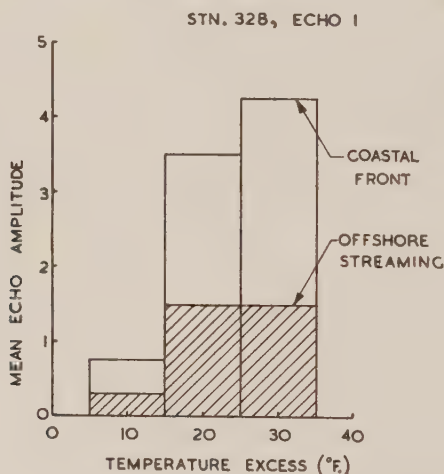


Fig. 12.—Comparison of mean echo strengths due to coastal fronts and offshore streaming. (Relative echo voltages  $\geq 10$  have been placed = 10. Results from 39 suitable days have been used.)

responsible, for example, for most of the superrefraction occurring on the New South Wales coast in winter, when temperatures are seldom high enough for offshore advection of warm air to be effective.

With a clear sky and the air nearly calm, nocturnal cooling of the ground results in a shallow surface layer of cold air, topped by a "radiation inversion". Provided the original wind allows an inversion to form, the surface wind will decrease as cooling proceeds, because the increasing stability of the lower layers retards the downward transfer of momentum through turbulence. In regions below a mountain slope the surface wind will increase later owing to the arrival of cold air that has drained down from the mountain. Air cooled over the land can be carried across the coastline by the residuum of the offshore general wind, by a katabatic wind, or more often by a combination of the two. Such a situation



is sometimes found to give rise to moderate superrefraction on seaward paths but the effect extends only a few miles from the coast, as the lower air is soon modified in passing over the warmer sea. The scale of the phenomenon has been confirmed by low-level aircraft flights carried out by M. Iliffe of this Laboratory (unpublished data) and from the analysis of data on ship echoes.

The maximum of superrefraction due to simple offshore streaming of warm air occurs in the early part of the night. A second maximum is frequently found to occur in the early morning, presumably owing to the supply of nocturnally-cooled air to the lower layers over the sea. The latter effect often seems to extend the duration of superrefraction of the offshore streaming type, which usually continues longer into the night than might be expected from the advection of warm air alone.

(iv) *Comparison of the Three Main Types.*—The three types of seaward superrefraction described above are distinguished by the different sources of the relatively cool underlying air. This may be provided by horizontal transport of sea-breeze air, by modification of warm air streaming over a cool sea, or by advection of nocturnally-cooled air.

These three processes are responsible for the three types of diurnal variation of sea- and shore-path echoes, namely :

Coastal front, sea breeze	..	..	..	1300–1900 hours
Offshore streaming	..	..	..	1700–0100 hours
Advection of nocturnally-cooled air	..	..	..	2400–0900 hours

In this list the three processes have been set out in decreasing order of the temperature excess required to produce them. The strength and seaward extent of the resulting superrefraction are in the same order.

The effect of a coastal circulation below the offshore stream on the strength of superrefraction has been tested by deriving the mean strength of a sea-path echo at station 328 on the north-west coast as a function of the temperature excess for the two types of situation. (The island producing this echo was 100 miles from the station and twenty-three miles from the nearest point of the mainland.) As shown in Figure 12 the mean echo strength with a coastal front was greater than with simple offshore streaming, for a given temperature excess. The temperature ranges of the two situations overlap, because the temperature excess at which a coastal circulation can be set up varies with the strength and direction of the general wind.

(v) *Minor Causes.*—Some of the observed cases of superrefraction are associated with two minor causes, cold fronts and subsidence. In a typical polar front, cool and relatively moist southern maritime air undercuts warm, dry, tropical continental air, but only in a few cases was the refractive index discontinuity at the front found to be sufficient to produce superrefraction.

Superrefraction due to downward bending at a subsidence inversion is not very important in Australia, the inversions usually being too high. However, some cases of superrefraction in south-west Australia and eastern New Guinea are believed to be due to subsidence at relatively low levels. Anticyclonic curvature of the wind field is sometimes particularly large in these regions.

(vi) *Land-path Echoes*.—Superrefraction on land-path echoes occurs almost exclusively at night-time and is closely related to radiation inversions.

Under relatively dry conditions, the diurnal curve shows a maximum near sunrise when the inversion is most intense. When dew forms, the maximum is found about midnight, with a subsequent dip and a later rise about sunrise. Each type of curve is illustrated in Figure 8. Examination of the monthly curves of diurnal variation for land-path echoes in north-west Australia has shown that the occurrence of a double maximum correlates more closely with months of high rainfall (i.e. wet soil) than with high humidity.

The temperature inversion alone can produce superrefraction, but when the ground surface is moist, evaporation causes an accumulation of water vapour near the surface, thus increasing the effect due to temperature. The double maximum may arise as follows. As cooling proceeds, superrefraction will intensify, unless or until the air temperature falls below its dew point. Water will then condense out of the lowest layers, reducing and even reversing the humidity gradient, and consequently decreasing the refractive index gradient. Immediately after sunrise the dew will be re-evaporated into the lower layers, causing a temporary increase in superrefraction until the inversion is finally destroyed.

While radiation inversions are the outstanding cause of superrefraction on land-path echoes, the minor causes, cold fronts and subsidence inversions, also affect echoes of this type when they occur.

#### V. VERY LONG-RANGE ECHOES

The most intense superrefraction is found off the north-west coast, where the trade wind stream carries hot, dry air over the coast most of the year. Some striking examples of very long ranges have been reported from this region, such as the detection of the IFF\* signal from a Ceylon-bound aircraft by station 47 (Geraldton) up to 1023 miles on February 14, 1944, and the North-West Cape report of echoes from Java (900 miles away) during September and October 1944. Regular observations were made for nearly a year by three Royal Australian Air Force radar stations in the northern part of this region, on echoes at ranges of 300 to 400 miles. The propagation paths were over sea, to Timor, and between Darwin and the Kimberleys (see map, Fig. 13). The detailed results are reproduced in Figure 13, in which the station locations and propagation paths are also indicated. Over the whole period the echoes were seen for 4 per cent. of the time. The only clear seasonal effect was a drop-out during December and January, during which months the north-west monsoon is most strongly developed. The echoes occurred most frequently during the night and early morning.

An outstanding feature of these observations is a marked tendency for several very long-range echoes to occur simultaneously. This indicates that they are affected by a single mechanism operating over a large area. The

\* IFF (Identification of Friend or Foe) comprises a transmitter triggered by the radar signal, giving a time-coded enhancement of the aircraft echo.

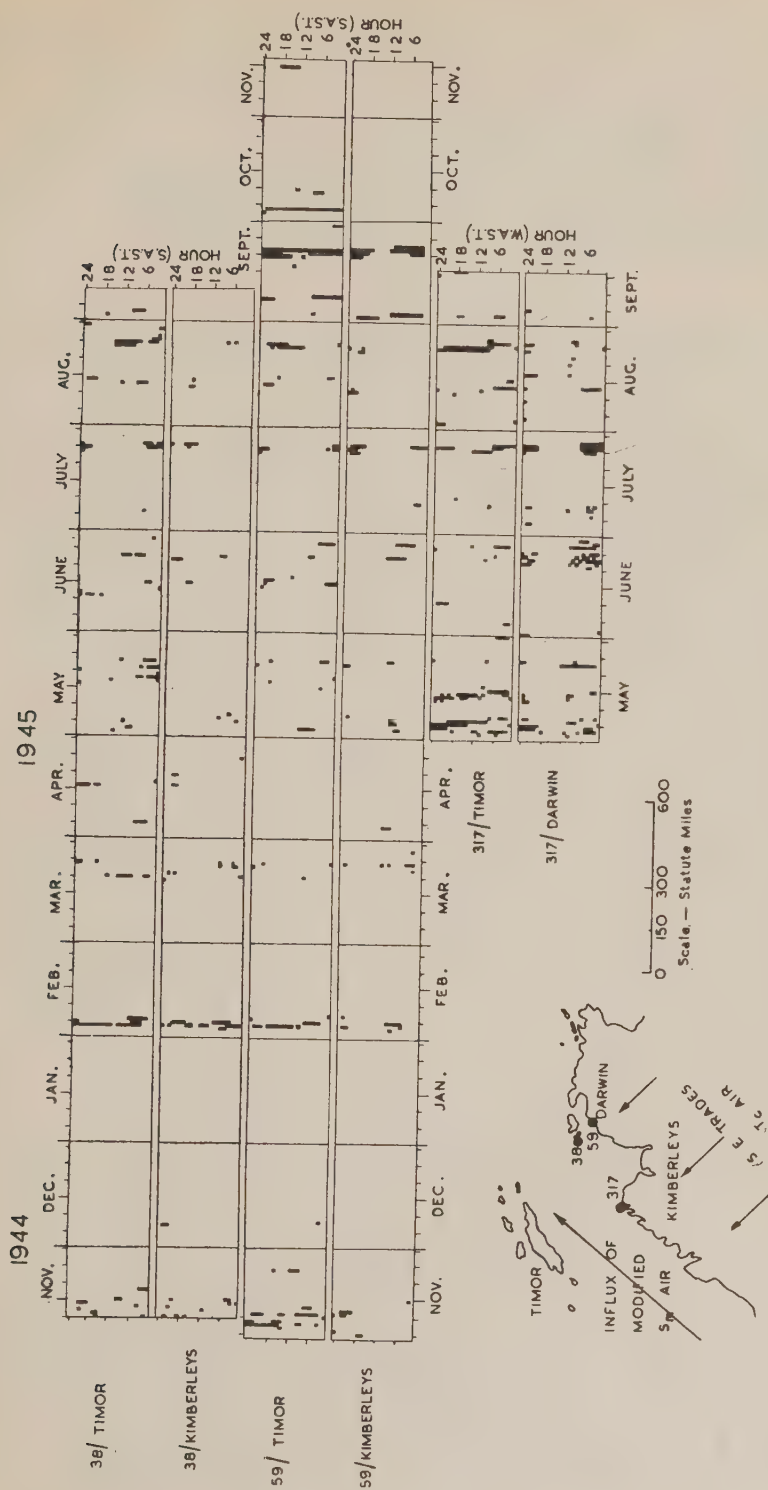


Fig. 13.—Occurrence diagram of very long-range echoes received by three stations (38, 59, 317) in northern Australia, November 1944–November 1945. The strip labelled “38/Timor” presents the hours in which station 38 received an echo from any portion of Timor. The map shows a typical situation associated with large-scale superrefraction.

characteristic behaviour was for the echoes to appear for a few hours or sometimes days at a time, with long intervals between. Many of these instances appeared to be associated with the influx of modified southern maritime air into the Timor Sea region, after sweeping round the west coast of the continent behind an eastward-moving polar front (see map, Fig. 13). The trajectory of this air is mainly over sea. Upon reaching the Timor Sea, this moist air is overlain by the dry tropical continental air streaming offshore. This situation is conducive to superrefraction.

## VI. REGIONAL SURVEYS

In this section the results will be correlated according to geographical region. Some degree of repetition is unavoidable.

(i) *Southern Australia*.—The mainland coast from Geraldton (Western Australia) to Brisbane (south Queensland) is considered here.\* Throughout the region there is a summer maximum, with the coastal front the principal cause of superrefraction. The weak superrefraction occurring in the winter is due mainly to nocturnally-cooled air carried out to sea, with simple advection operating in the intermediate seasons. In addition, fleeting effects are caused by passing cold fronts, with subsidence having some influence in the south-west corner. The superrefraction associated with coastal fronts, offshore streaming, and radiation inversions is most likely to occur in the rear of anticyclones, where the gradient wind is usually offshore.

(ii) *North-west Coast*.—Superrefraction is strongest on the north-west coast, where echoes from very long ranges are sometimes reported. Observations on echoes at moderate ranges (80–150 miles) show pronounced diurnal effects masking the influence of synoptic variations, which are generally weak in tropical areas. The coastal front dominates the radio-meteorology of this region, commencing in early spring, increasing in height and intensity as summer progresses, and falling away again in autumn. Superrefraction shows maxima in spring and autumn, decreasing in summer when the front is usually too high (5000–6000 ft.) for adequate guiding of 200 Mc/s. waves by the resulting duct. In the winter, moderate superrefraction occurs owing to advection ducts.

(iii) *Northern Australia*.—The coastal front of the north-west coast seldom extends beyond Broome. Sea-breeze effects are important, however, at the end of the dry season (September–November) before the north-west monsoon sets in. The trade-wind stream carries continental air over the coast throughout the year except in the wet season (December–February). Seaward superrefraction to moderate distances has its maximum in July–August, when the offshore stream is most strongly established, and occurs but rarely in the wet season.

Important day-to-day variations occur and it was thought that they might be simply related to afternoon temperature and humidity about 100 miles inland, in the seasons when superrefraction is due to advection of warm air over the sea.

\* No stations were operating on the coast of the Great Australian Bight or Tasmania. However, the statements made in this subsection are believed to apply to the Bight coast.



A statistical investigation of superrefraction at Darwin showed, however, little correlation with conditions at an inland site, except that superrefraction was absent when the air was excessively cool. The complex nature of the radio-meteorology of the region was exemplified by the results of the Radiophysics Laboratory expedition to Darwin in 1944. The conditions governing moderate-distance superrefraction are more localized in this region than in southern Australia, as is shown by the much smaller correlation in day-to-day variations between stations 50–100 miles apart. Local topography apparently plays a large part.

Land-path echoes at moderate distances in this region show clearly the characteristic nocturnal maximum of radiation inversions. The occurrence of superrefraction is greatest in the early part of the dry season, when moist vegetation is abundant. A rough inverse correlation with cloud cover is found, but day-to-day variations do not correlate closely with the diurnal range of temperature at a single inland station. Conditions over the actual propagation path would need to be known, since radiation phenomena show large local variations.

(iv) *North-east Australia*.—Superrefraction on the Queensland coast is weak and infrequent compared with the rest of Australia. The south-east trade wind is the dominant meteorological factor in this region, bringing in a deep stream of tropical Pacific air around an anticyclone in the Tasman Sea. This occurs throughout the year except for short periods when the prevailing pattern is disturbed by a polar front or tropical cyclone. Nearly all the superrefraction on this coast is associated with the approach of polar fronts, the maximum occurrence being found one to two days before the arrival of the front. The flow is usually offshore ahead of a front, the air being diverted trade-wind air, modified over the continent. In consequence of the north-west to south-east orientation of the typical polar front in this region, the overland trajectory is greatest some distance ahead of the front. Frontal passages and hence superrefraction occur at longer intervals here than in the south. The seasonal maximum occurs in October–December when the air over the coast is warmest and driest relative to the sea.

(v) *New Guinea*.—Superrefraction occurs in New Guinea only to a limited extent. Available information is restricted almost entirely to the eastern half of the island, but it is believed that very little, if any, superrefraction occurs anywhere on the south coast, or on the central stretch of the north coast.

American results(7) have confirmed that surface ducts are set up in the Geelvink Bay area by the “wam brow”, a local föhn wind. This wind occurs for periods of four to eight days during the latter part of the south-east season (July–October), when an unusually intense anticyclone is established over the Australian continent(8, 9).

On the central-northern coast the only regular observations available are from Aitape, where superrefraction is absent. The behaviour is probably similar in the whole of this region and on the south coast also, though there may be slight effects related to the sea breeze.



Superrefraction is more marked in and north-west of the Vitiaz Straits, local sea-breeze effects being responsible. The echoes occur almost entirely between 1000 and 2000 hours. (Sea breezes in general commence earlier in New Guinea than in northern Australia.) The seasonal variation shows a maximum in August–November, which is the period of maximum sea-breeze influence according to the surface wind statistics at Madang. Superrefraction in this region exhibits large local variations, echo occurrences at neighbouring stations showing little correlation on a day-to-day basis.\*

Larger-scale superrefraction of moderate intensity occurs rather infrequently in the Goodenough Island–Milne Bay region, where results from stations 100 miles apart correlate well. All of the principal occurrences (and most of the minor ones) are associated with a cold front lying north-west to south-east along the Queensland coast (see map, Fig. 14). It appears that the superrefraction is

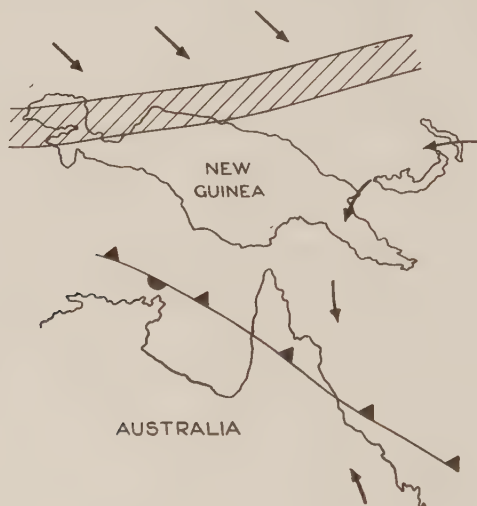


Fig. 14.—Typical situation associated with superrefraction in eastern New Guinea, showing curvature of wind-field between intertropical convergence zone and trailing polar front.

related to the large anticyclonic curvature of the wind field, with consequent subsidence, which is found in such a situation between the trailing cold front and the Intertropical Convergence Zone. The echo variations do not correlate with surface meteorological conditions nor is the superrefraction associated with the cold front itself, as it occurs when the front is 500–1000 miles from the area. The seasonal behaviour shows a maximum of superrefraction occurrence in December–March. The diurnal variation is not marked, but there is a moderate maximum in daylight hours.

In the Rabaul area of New Britain, the south-east trades blow for most of the year. Available information covers only a few months, but weak superrefraction is found to occur when the wind blows from the north-east quadrant,

approaching Rabaul from New Ireland. Under these conditions, some heating will occur in crossing New Ireland, together with a föhn effect due to the high mountains.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R., in collaboration with the Royal Australian Air Force. It was initiated by Dr. D. F. Martyn, while a member of the Australian Army Operational Research Group, and P. Squires. Responsibility for the analysis later passed to the author. Assistance in the statistical and physical analyses was given by J. W. Reed, J. K. Strachan, and Miss B. Lippmann. J. W. Reed also visited the stations in the selected list to maintain liaison with the observers. Acknowledgments are due to the radar personnel of the R.A.A.F. for making the observations, to the Directorate of Meteorological Services, R.A.A.F., for the supply of meteorological data, to the Women's Australian Auxiliary Air Force personnel who assisted in the statistical work, and to the Australian Military Forces for undertaking similar observations at some of their radar stations.

## VIII. REFERENCES

- (1) BOOKER, H. G.—Telecommunications Research Establishment Report No. T.1820 (1945).
- (2) BOOKER, H. G.—*J. Instn. Elect. Engrs.* IIIA, **93**: 69-78 (1946).
- (3) MARTYN, D. F., and SQUIRES, P.—Australian Radio Propagation Committee, Ionospheric Bulletin for October 1944: 5-9.
- (4) CASSIDY, M.—R.A.A.F. Meteorological Services, Melbourne, Weather Development and Research Bulletin (Southern Australian Region) No. 1: 5-28 (1945).
- (5) HATCHER, R. W., and SAWYER, J. S.—*Quart. J. R. Met. Soc.* **73**: 391-406 (1947).
- (6) CRAIG, R. A., KATZ, I., and HARNEY, P. J.—*Bull. Amer. Met. Soc.* **26**: 405-10 (1945).
- (7) ANDERSON, P. A., FITZSIMMONS, K. E., GROVES, G. M., and STEPHENSON, S. T.—N.D.R.C. Report No. CP-30 (1945).
- (8) HOGAN, J.—R.A.A.F. Meteorological Services, Melbourne, Tropical Research Bulletin, No. 3: 3-6 (1944).
- (9) BOND, H.—R.A.A.F. Meteorological Services, Melbourne, Tropical Weather Research Bulletin, No. 6: 3-11 (1944).

# A NEW INTERFEROMETER

By E. R. JOHNSON\* and J. F. M. SCHOLES\*

(Plate 1)

[*Manuscript received September 28, 1948*]

## *Summary*

A new interferometer, especially suitable for studying aerodynamic fields, in which a parallel beam of light passes once through the field, is described. The dividing plates and the optical components in the auxiliary beam are small, though the main beam covers a wide field. The relevant theory and an outline of a method of adjustment are given; and possible development to replace lenses with mirrors and to improve light transmission is indicated.

## I. INTRODUCTION

The application of interferometric methods to the study of aerodynamic fields gives a complete quantitative picture of the field without disturbing the flow by the introduction of probes. This is specially important in transonic and supersonic fields where probes introduce shock waves, seriously altering the flow pattern. It is usually possible to determine the density at some point in the field from pressure and temperature measurements elsewhere. The interferometer measures the density throughout the field relative to this point. From this knowledge of the density field, use of the principles of conservation of mass, momentum, and total energy enables the determination of all physical quantities throughout the field.

The sensitivity of the interferometer is such that satisfactory accuracy is obtained with aerodynamic apparatus of moderate size. The ability to take a series of extremely short duration photographs covering the whole field makes possible the study of unsteady processes in detail. Only in recent years has the interferometer been applied extensively to the study of aerodynamic fields, and the arrangement used has been almost exclusively that of Mach(1) and Zehnder(2).

The fields to be covered are considerably larger than in other applications of the interferometer. For ease of reduction of the results in two-dimensional and axially-symmetrical flow fields the beam traversing the field should be collimated. To avoid doubling of the image by refraction the beam should pass only once through the field, and in practice, the auxiliary beam must be well removed from the main beam. White light fringes are necessary to identify fringes that are dislocated across density discontinuities. The Mach-Zehnder

\* Division of Aeronautics, C.S.I.R.

interferometer is the only one that satisfies these conditions. This, however, requires two plane parallel dividing plates of width  $\sqrt{2}$  times the field of view, and to obtain these of the required size is both difficult and expensive.

The interferometer to be described, by collimating the beams after they pass through the dividing plate, makes it possible to use a small dividing plate. This involves the use of two pairs of lenses instead of one, but the pair in the auxiliary beam can be of small aperture, and the large lenses in the main beam may be replaced by parabolic mirrors. It is necessary to introduce compensating plates in order to make the optical path lengths equal in the two pairs of lenses.

## II. THE INTERFEROMETER

A diagram of the optical system of the interferometer is shown in Figure 1. Light from the source  $S$  is focused by the condenser  $C$  on to a pinhole  $P$ . This pinhole is at the common focus of lenses  $L_2$  and  $L_1$  as imaged in the semi-reflecting dividing plate  $D$ . The beam is divided into two parts by  $D$  and these are collimated by lenses  $L_1$  and  $L_2$ . After reflexion at  $M_1$ , beam I is focused by  $L_1'$  at  $P_1'$ . Beam II is reflected at  $M_2$  and focused by  $L_2'$  after reflexion at the semi-reflecting plate  $D'$  on a point  $P_2'$  adjacent to  $P_1'$  ( $P'$ ). The two beams are thus reunited by  $D'$  and two coherent images of  $P$  are formed at  $P'$ , giving rise to interference fringes on the photographic plate. The camera lens focuses the field  $F$  on the plate.

## III. THEORY

The Michelson, the Mach-Zehnder, and this interferometer are two-beam interferometers, and in a theoretical discussion of these, Hansen(3) arrives at the conclusion that the fringes are localized where rays, coming from the same point of the dividing plate and having passed through the same distance by different paths within the interferometer, intersect. Schardin(4) shows that this is the plane of intersection of the perpendicular bisectors of lines joining corresponding points in the two virtual sources formed by imaging the actual linear source in the mirrors.

The fringe pattern formed on the plate is in every respect the same as that arising from two coherent sources at the images  $P_1'$  and  $P_2'$  of  $P$ . If the focal lengths of lenses  $L_1$  and  $L_1'$  and of  $L_2$  and  $L_2'$  are equal, and the beams are properly collimated, images  $P_1'$  and  $P_2'$  are both equal in size to  $P$ . Then by the Lagrange invariant convergence angles  $U_1'$  and  $U_2'$  are equal, since these rays are two portions of the same ray incident on  $D$ . If the focal lengths of lenses  $L_1$  and  $L_1'$  are nearly but not exactly equal, the beam I between them can be accurately collimated, but the beam II need not be collimated. There are then two degrees of freedom, axial movements of  $L_2$  and of  $L_2'$  (which, in general, have slightly different focal lengths) to form the image  $P_2'$  in the same plane and of the same size as  $P_1'$ , satisfying the convergence conditions.

If a fringe system of parallel lines is set up by rotating  $M_2$  through an angle  $\beta$  and  $D'$  through  $\alpha$  from their symmetrical positions so that the rays of beams I and II through an axial point  $P$  intersect on the plate, the fringe spacing is



$b = \lambda \cdot P_1'X/P_1'P_2'$  where  $\lambda$  is the wavelength of the light used. By reasoning, similar to that of Schardin (p. 425), it can be shown that  $b = \frac{1}{2}\lambda/(\alpha - \beta)$ .

In order to allow for a finite aperture at  $P$ , consider the plane  $f$  which is conjugate to  $F$  in lens  $L_1$  (see Fig. 1) and therefore conjugate to the plate  $XY$

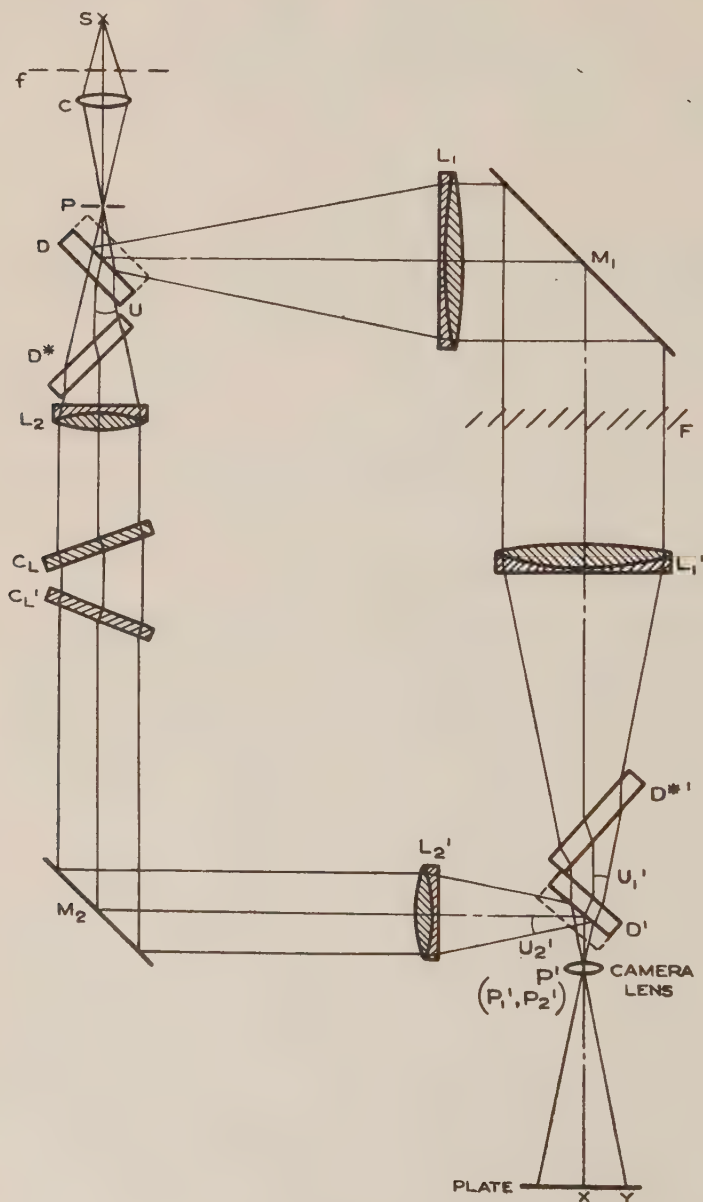


Fig. 1.—Experimental arrangement.

in the system  $L_1L_1'$ . In general, the planes  $f$  and  $XY$  will not be conjugate in the system  $L_2L_2'$  and this will cause blurring of the fringes. Though this is



tolerable for a small aperture, it is advantageous to design the auxiliary system to meet this condition and the theory is then the same as that given by Schardin for an interferometer with external collimation.

### (a) Compensation for White Light Fringes

In general the path lengths within the lenses  $L_1$ ,  $L_1'$  and  $L_2$ ,  $L_2'$  are different, and theoretically it is necessary to insert in the collimated beam II a plate of each of the glasses forming the doublet lenses to make the optical path lengths equal for all colours. In practice, it is possible to choose a glass, intermediate between the glasses used in the lenses, that gives sufficient compensation.\* This plate is divided into two of equal thickness whose combined thickness is slightly less than that needed to compensate the lenses. By equal rotations in opposite directions of these plates about parallel axes, compensation can be achieved without lateral shift or dispersion of the beam (see Fig. 1,  $C_L$  and  $C_L'$ ).

Though the equal dividing plates  $D$  and  $D'$  compensate each other as far as path lengths are concerned, the white light fringes are those shown in Plate 1, Figure 1. The plates introduce lateral chromatic aberration of equal magnitude, but of opposite sign, into the beams. Thus if the fringes in green light run parallel to the plane of the paper in Figure 1, so that the line  $P_{1G}'P_{2G}'$  is perpendicular to the paper at  $P'$ , then for blue light  $P_{1B}'$  is to the left of  $P_{1G}'$  and  $P_{2B}'$  is to the right of  $P_{2G}'$ . The blue fringes are no longer parallel to the paper; the fringe which passes through  $X$ , rises above  $Y$  on the plate, and vice versa for red light. The angles between sets of parallel fringes in different colours are small, so that dark fringes in all colours coincide along a line inclined to the fringes as shown in Plate 1, Figure 1. Here the fringes are approximately horizontal but the line of coincidences is inclined at about  $45^\circ$  to the horizontal. Though it does not show in the photograph, the fringes continue beyond the short dark portions in a series of coloured strips.

This could be corrected by introducing compensating plates alongside  $D$  and  $D'$  as shown by dotted lines in Figure 1. A better arrangement is to introduce the plates at  $D^*$  and  $D^{*}$ . The lateral dispersion is then rectified outside the lens systems and rays in all colours traverse the same path through the lenses. (See also the next section.)

### (b) Effect of Aberrations

Since the object  $P$  is limited to a small axial one, only spherical aberration of the lenses, and spherical aberration, coma, and astigmatism, introduced by the oblique parallel plates  $D$  etc. are present in monochromatic light.

The lenses are figured to remove spherical aberration. The plates  $D$  etc. are made identical in manufacture†; so the spherical aberration introduced by  $D$  and  $D^*$  in beam II equals that introduced by  $D'$  and  $D^{*}$  in beam I. It is

\* Plate 1, Figure 2, was obtained using "Crystallex" plate glass for compensation.

†  $D$  should be identical with  $D'$  and  $D^*$  with  $D^{*}$ , but these pairs need not be identical. The limit to their difference in thickness is set by the need to correct lateral dispersion (see previous section).

sufficiently accurate to say that the longitudinal magnification of the system  $L_2L_2'$  is unity and consequently the spherical aberration in the recombined beams is equal and the fringe pattern is not affected.

The lateral dislocation of the sagittal rays by the plates  $D$  and  $D'$  is rectified outside the lens systems by the plates  $D^*$  and  $D^{*'}.$  The astigmatism and chromatic aberrations introduced by the oblique parallel plates are the same in both beams and do not affect the fringe pattern.

The effect of chromatic aberration in the lenses is shown in Plate 1, Figure 2. If  $P_1'$  and  $P_2'$  fall on a plane perpendicular to the axis, the fringes approximate to a set of parallel straight lines. If  $P_1'$  and  $P_2'$  fall one behind the other on the axis, the fringe pattern is a series of concentric circles. In intermediate positions the fringes are a series of arcs, being the lines of intersection of a plane with a set of hyperboloids of revolution. If  $P_{1G}'$  and  $P_{2G}'$  are made to fall on a plane perpendicular to the axis so that the fringes in green light are a set of parallel straight lines, then, because of the different longitudinal chromatic aberration of the two pairs of lenses,  $P_1'$  and  $P_2'$  for other colours will not fall in a plane perpendicular to the axis. The fringes in these colours will be a series of arcs which, on superposition in white light, give rise to the pattern shown in Plate 1, Figure 2. There is the usual set of coloured fringes, not visible in the photograph, on either side of the dark central fringes.\* This curved pattern is not a serious disadvantage since it is a small effect for which allowance can be made in interpreting the photographs.

#### IV. ADJUSTMENT

The adjustment of the interferometer is most easily done by first setting up a Mach-Zehnder interferometer with the plates  $D$  and  $D'$  and the mirrors  $M_1$  and  $M_2$  by any of the well-known methods. This is adjusted to exhibit one focused zero-order fringe over the whole field.  $D^*$  and  $D^{*'}.$  are set up so that the zero-order fringe is retained. Lenses  $L_1$  and  $L_2$  are adjusted so that the beams entering them correspond at the plate  $D$ , and  $L_1', L_2', C_L$  and  $C_L'$  are placed so as to receive the full collimated beams from  $L_1$  and  $L_2$ . The fringes which appear in the field are focused on, and perpendicular to, the slit of a spectroscope.

Using white light, the spectrum is crossed by a set of fringes. Rotation of  $C_L$  and  $C_L'$  by equal and opposite amounts brings into the spectrum one fringe perpendicular to the slit. This fringe, in general, is not straight because of inexact compensation. By varying the angle of rotation of  $C_L$  and  $C_L',$  and altering the path difference with  $D',$  any desired portion of the spectrum may be brought to a state of best compensation. The interferometer is then adjusted by movement of  $L_1'$  and  $L_2'$  so that the zero-order fringe now visible is of infinite width. The fringes are adjusted to the desired pattern by rotation of  $M_2$  and  $D'.$

The pairs of lenses  $L_1$  and  $L_1',$  and  $L_2$  and  $L_2',$  were made to the same specification and thus differed only slightly in focal length. Because of the limitation

\* As noted before, the photographs in Plate 1 were taken using "Crystallex" plate glass for the compensating plates. This gives rise to the slight curvature of the dark fringes in the photograph. The effect explained above causes the curved locus of points of coincidence of fringes in all colours.

of the size of the pinhole by the blurring of the fringes discussed in Section III, it has not been found necessary to alter the system  $L_2, L_2'$  after it is set up as described above.

## V. ILLUMINATION

In this system there is a greater loss of light than there is in the Mach-Zehnder interferometer. For much aerodynamic work it is desirable to keep the exposure time down to the order of 1-2 microseconds to stop unsteady flow. This requires a source of very high intrinsic brightness, conflicting with the requirement of a monochromatic one.

Assuming that (i) at each air-glass surface there is a loss of 4 per cent., (ii) that the aluminized mirrors are 85 per cent. efficient, and (iii) that 30 per cent. is absorbed in the semi-reflecting films, the amount of light reaching the plate is approximately 12 per cent. of that incident on the pinhole. A Wratten No. 34 filter reduces this to about  $2\frac{1}{4}$  per cent.

The photographs (see Plate 1, Figs. 3 and 4) were obtained with an exposure of 1-2 microseconds, using as a source a Mullard LSD2 discharge tube through which a 2 microfarad condenser charged to 10 kilovolt was discharged. A Wratten No. 34 filter was used. The pinhole was  $\frac{1}{4}$  mm. in diameter; the focal length of the lenses  $L_1$  and  $L_2'$  was 510 mm., and of  $L_2$  and  $L_2'$  was 300 mm.; their relative aperture was  $f/12$ .

By treatment of the air-glass surfaces to reduce the reflexion loss and using semi-reflecting films of higher efficiency, the amount of light reaching the plate could be increased two- or threefold.

If the system  $L_2L_2'$  be altered to give two more degrees of freedom, so that the image of  $f$  in beam II falls on the plate at the same magnification as that of beam I, the size of the pinhole can be increased to give a greater intensity on the plate.

## VI. FUTURE DEVELOPMENT

It is proposed to replace the large lenses with parabolic mirrors in the arrangement shown in Figure 2.\* The main beam, after passing through a small hole in the centre of the large plane mirror, is collimated by the parabolic mirror and reflected through the disturbance by the large plane mirror. After reflexion by the second large plane mirror and parabolic mirror, it is recombined at the second semi-reflecting plate with the auxiliary beam which has been separated from the main beam at the first semi-reflecting plate and has traversed the path  $M_1$ , lens,  $M_2 \dots M_7$ , lens,  $M_8$ .

## VII. CONCLUSION

A modification of the Mach-Zehnder interferometer, for use in aerodynamic research, which covers a large field without the need for large plane parallel dividing plates, has been tested and shown to be practicable.

\* Note.—This does not include the necessary compensating plates. The tunnel windows will contribute to the compensation.

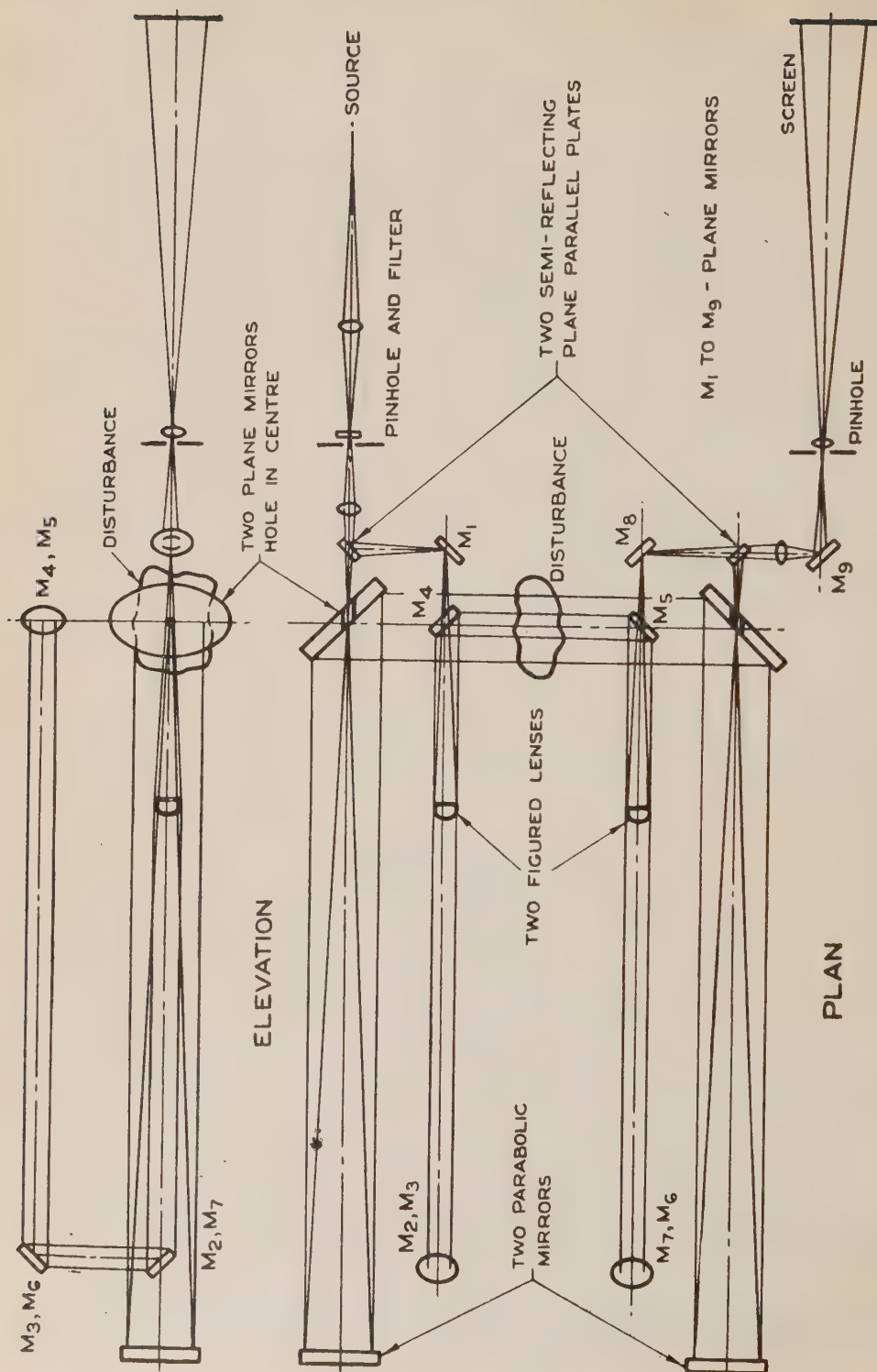


Fig. 2.—Proposed interferometer.





Fig. 1



Fig. 2

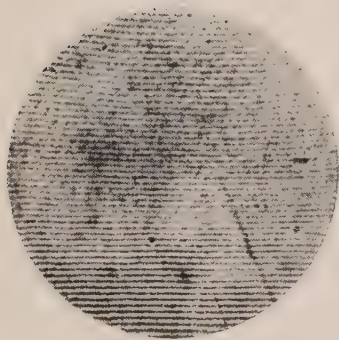


Fig. 3

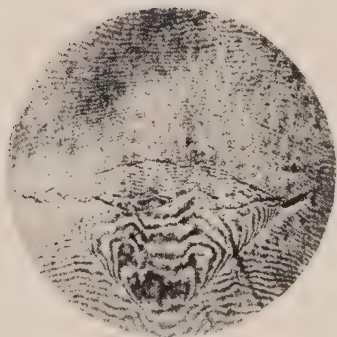
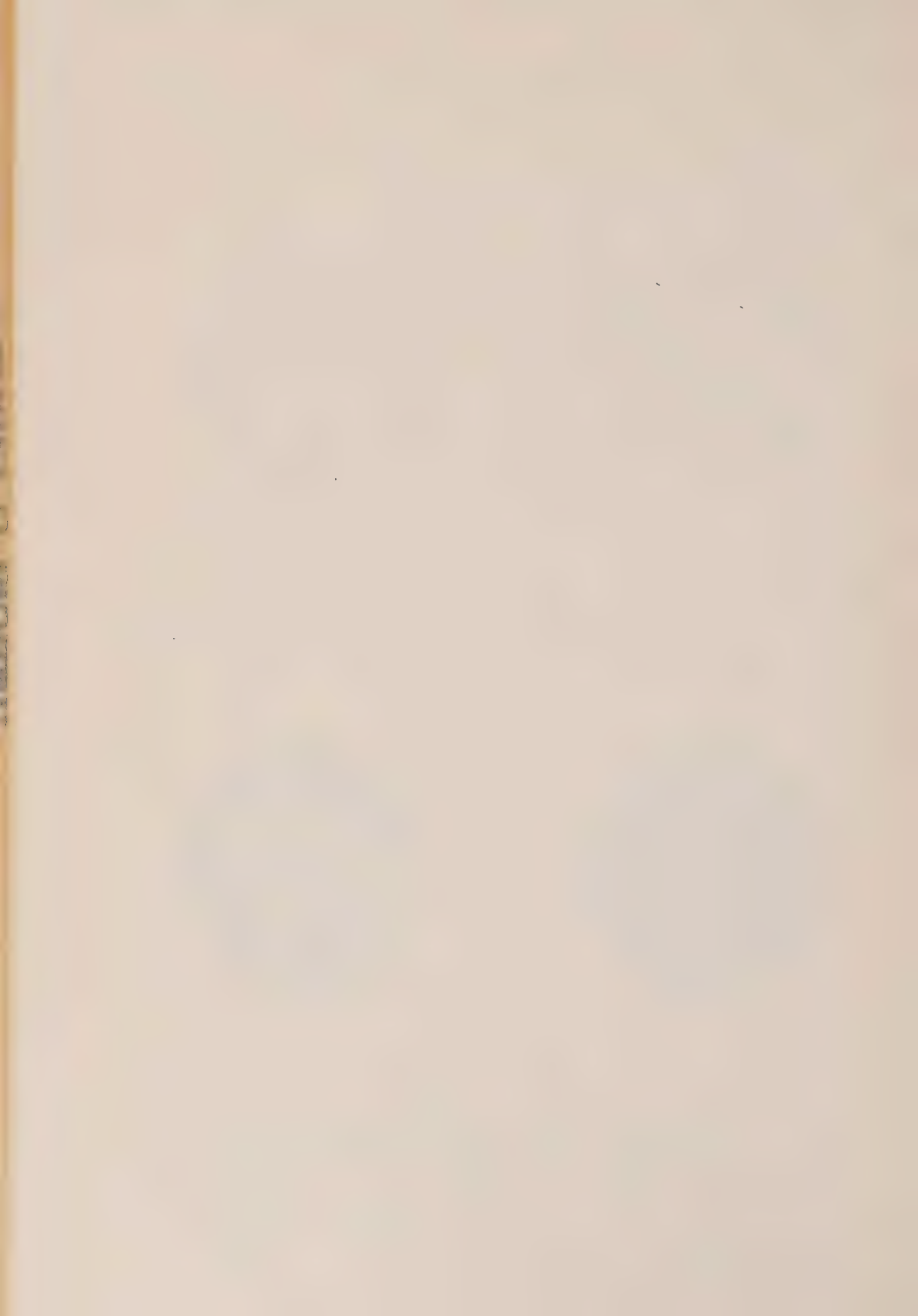


Fig. 4





## VIII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Aeronautics, C.S.I.R.

Our thanks are due to Mr. E. N. Waterworth of Hobart for his cooperation in the production of the necessary optical components, and to Messrs. W. H. Steel and J. W. Pearce of the Division of Physics, C.S.I.R., for helpful criticism of this paper before it appeared as a Divisional Report.

## IX. REFERENCES

- (1) MACH, L.—*S.B. Akad. Wies. Wien.* (2A) **101**: 107 (1892).
- (2) ZEHNDER, L.—*Z. InstrumKde* **11**: 275 (1891).
- (3) HANSEN, G.—*Z. Tech. Phys.* **12**: 436 (1931).
- (4) SCHARDIN, H.—*Z. InstrumKde* **53**: 396, 424 (1933).

## EXPLANATION OF PLATE 1

- Fig. 1.—White light fringes without plates  $D^*$  and  $D^{*'}.$
- Fig. 2.—White light fringes in compensated interferometer.
- Fig. 3.—Undisturbed field with LSD2 source and Wratten No. 34 filter.
- Fig. 4.—Flow pattern from a 25 mm. nozzle with LSD2 source and Wratten No. 34 filter.

# THE VISIBLE ABSORPTION SPECTRUM OF BROMINE IN SOLUTION

By N. S. BAYLISS,\* A. R. H. COLE,\* and B. G. GREEN\*

[*Manuscript received August 12, 1948*]

## *Summary*

The visible absorption spectrum of bromine was measured in *n*-hexane (a normal solvent) and in concentrated sulphuric acid (at room temperature and at  $-70^{\circ}\text{C}.$ ), phosphoric acid, and ethanol (at  $-70^{\circ}\text{C}.$ ) as associated solvents. The absorption continuum at 4150 Å. in the gas is displaced slightly to the red in *n*-hexane although other normal solvents may displace it to the violet. The displacement to the violet is greater in associated solvents, ranging from  $650\text{ cm.}^{-1}$  in sulphuric acid to over  $2000\text{ cm.}^{-1}$  in ethanol. The displacement in sulphuric acid glass at  $-70^{\circ}\text{C}.$  is nearly double the value at  $18^{\circ}\text{C}.$  The displacements in associated solvents are explained in terms of the effect of the semi-rigid cage of solvent molecules surrounding each bromine molecule, and the assumption of "solvates" is unnecessary and undesirable. The spectra in solution are between 40 and 70 per cent. more intense than in the gas, about double the increase that would be predicted by Chako's treatment of the effect of Lorentz-Lorenz forces. The differences in the shape and  $\epsilon_{\text{max}}$  between solutions in sulphuric acid at  $18^{\circ}\text{C}.$  and at  $-70^{\circ}\text{C}.$  are due to the different statistical distribution of bromine molecules between the vibrational energy levels.

## I. INTRODUCTION

In spite of the satisfactory state of the theory of the absorption spectra of diatomic gases, the changes in the position and intensity of the absorption bands of even simple molecules under the influence of solvents are still far from complete interpretation. The work presented in this paper is in continuation of a programme(1) to use simple molecules such as bromine, whose gaseous absorption spectrum is reasonably well understood, as the experimental material in an attempt to isolate the factors that are responsible for solvent effects on spectra.

Previous investigators of the absorption spectrum of bromine in solution include the following: Bovis(2), Gillam and Morton(3), Child and Walker(4), Aickin, Bayliss, and Rees(1), and Prikhotko(5). Although they are not in perfect agreement, some general conclusions may be drawn from their results. The absorption continuum responsible for the visible colour of bromine gas ( $\epsilon_{\text{max}}=152$  at 4150 Å.) is increased somewhat in intensity in solution, and the position of the maximum is displaced slightly, usually towards higher frequencies. In "normal" or non-associated solvents the displacement is of the order of a few hundred  $\text{cm.}^{-1}$ ; in associated solvents it is rather greater—of the order of  $1000\text{ cm.}^{-1}$  to higher frequencies. At low temperatures the displacement appears

\* Department of Chemistry, University of Western Australia.

to be greater, amounting to  $1300\text{ cm}^{-1}$  in pentane at liquid hydrogen temperatures(5).

This paper describes measurements on the visible absorption spectrum of bromine in *n*-hexane as a typical normal solvent, and in ethanol, concentrated sulphuric acid, and concentrated phosphoric acid as associated solvents. Measurements in sulphuric acid are described at  $-70^{\circ}\text{C}$ . as well as at room temperature in order to study the effect of freezing the solvent to a rigid glass. The measurements in ethanol were carried out at  $-70^{\circ}\text{C}$ . in order to avoid the reaction between bromine and alcohol which is appreciable at room temperature.

## II. MATERIALS

*Bromine*.—A B.D.H. "Analar" sample was used which contained less than 0.0005 per cent. of iodine according to the Analar tests.

*n-Hexane*.—A B.D.H. sample was washed repeatedly with concentrated sulphuric acid and distilled. We could not remove a trace of benzene which could be detected by its absorption spectrum at 2600 Å.

*Ethanol*.—Benzene was removed from a commercial sample of absolute alcohol by adding some water and fractionating. The purified sample was completely transparent to 2200 Å.

*Sulphuric acid*.—B.D.H. "nitrogen-free" sulphuric acid (98 per cent.) was diluted with water to 90 per cent. This procedure was found necessary to void crystallization and to promote the formation of a glass when the acid was cooled to  $-70^{\circ}\text{C}$ .

*Phosphoric acid*.—The B.D.H. sample of orthophosphoric acid (sp. gr. 1.75) was transparent to 3200 Å.

## III. EXPERIMENTAL

Measurements at room temperature were carried out by the method of photographic spectrophotometry described by Aickin, Bayliss, and Rees(1). For the measurements at  $-70^{\circ}\text{C}$ ., the solutions were contained in a Hilger "Presil" cell, 4 cm. long, mounted in a thermally insulated copper bath containing a mixture of solid carbon dioxide and ethanol. The cell was held tightly between two re-entrant side tubes with outside silica windows, thus providing air gaps which protected the cell windows both from the bath liquid and also from deposition of frost from the atmosphere.

The results are expressed in terms of the molar extinction coefficient  $\epsilon$  defined by the usual formula  $\log_{10}(I_0/I) = \epsilon cd$ , where  $I_0$  and  $I$  are respectively the incident and transmitted intensities,  $c$  is the concentration of bromine in mole litre $^{-1}$ , and  $d$  is the length of the absorbing path in cm. The bromine concentration varied between 0.0002 and 0.0015 mole litre $^{-1}$ .

## IV. DISCUSSION

The results are shown graphically as absorption curves in Figure 1, and are summarized in Table 1, which for comparison includes data obtained by Aickin, Bayliss, and Rees(1) in other solvents. The absorption continuum under

consideration is known to consist of two components,  $A$  ( $\lambda_{max}$  in gas = 4150 Å.) and  $B$  ( $\lambda_{max}$  in gas = 4950 Å.), which have been identified with the electronic transitions— $A$ :  $^1\Pi_u \leftarrow ^1\Sigma_g^+$ ;  $B$ :  $^3\Pi_0^+ \leftarrow ^1\Sigma_g^+$ ,  $^3\Pi_{1u} \leftarrow ^1\Sigma_g^+$  (Acton, Aickin, and Bayliss, 6; Bayliss and Rees, 7; Rees, 8; however, see Mulliken, 9). The data in Table 1 refer to the stronger  $A$  component, and include the displacement ( $\Delta\nu_{max}$ ) of the absorption maximum in each solvent, and also the intensity of the continuum expressed in terms of the oscillator strength  $f$  as defined by the relation (10)

$$f = 4.31 \times 10^{-9} f \gamma \epsilon d \nu$$

$$\approx 4.31 \times 10^{-9} \gamma \epsilon_{max} \Delta\nu_{\frac{1}{2}} \dots \dots \dots (1)$$

The factor  $\gamma$  contains the refractive index of the solvent; but in view of the uncertainty about its true value it is taken as unity for the time being in accordance with the suggestion of Mulliken and Rieke (10). The half-width of

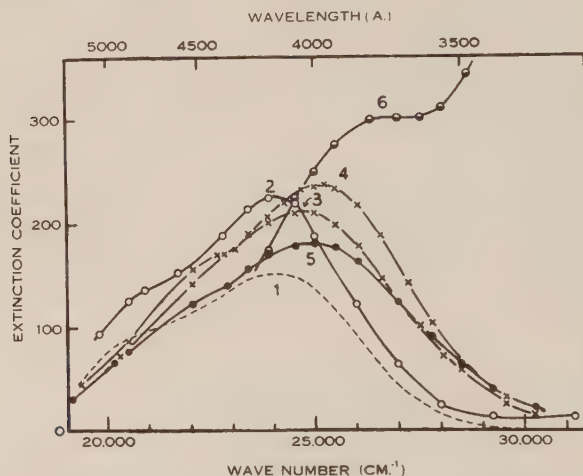


Fig. 1.—Visible absorption spectrum of bromine solutions showing molar extinction coefficient  $\epsilon$  plotted against wave number. The maxima are numbered as follows: 1,  $\text{Br}_2$  gas; 2, in  $n$ -hexane; 3, in sulphuric acid at 18° C.; 4, in sulphuric acid at -70° C.; 5, in phosphoric acid; 6, in ethanol at -70° C.

the continuum,  $\Delta\nu_{\frac{1}{2}}$ , is the separation in  $\text{cm}^{-1}$  between the points where  $\epsilon = \frac{1}{2}\epsilon_{max}$ . Owing to the contribution made by the  $B$  component on the low frequency side of the  $A$  maximum, we have assumed for the purpose of calculating  $\Delta\nu_{\frac{1}{2}}$  that the  $A$  component is symmetrical about  $\nu_{max}$ , and  $\Delta\nu_{\frac{1}{2}}$  has been taken as twice the separation between  $\nu_{max}$  and the point on the high frequency side where  $\epsilon = \frac{1}{2}\epsilon_{max}$ .

From Figure 1 and Table 1, the solvent effect is seen to consist in a displacement of the  $A$  maximum ranging from -100  $\text{cm}^{-1}$  ( $n$ -hexane) to 500  $\text{cm}^{-1}$  (carbon tetrachloride) in the normal or non-associated solvents, and from 650  $\text{cm}^{-1}$  (sulphuric acid) to 1400  $\text{cm}^{-1}$  (water) in the associated solvents. Cooling sulphuric acid to a glass at -70° C. nearly doubles the displacement. The increase in the intensity  $f$  varies from 40 per cent. in  $n$ -hexane to nearly



70 per cent. in sulphuric acid. The displacement and the intensity of the A component in ethanol are almost impossible to estimate with any accuracy owing to the proximity of a very strong absorption band in the ultraviolet. Our results in ethanol are at variance with those of Child and Walker(4) in that we obtain a considerably greater value for the apparent  $\epsilon_{max}$  at 3740 Å. and a much shallower minimum in the absorption curve before the onset of the strong ultraviolet absorption band. We believe the discrepancy to be due to the behaviour of the ultraviolet band, which is under investigation and will be discussed in a future paper.

(a) *The Effect of Solvents on  $\nu_{max}$*

The smaller displacement of  $\nu_{max}$  in non-associated solvents has been attributed to the operation of van der Waals' forces, and the larger displacement in associated solvents to the formation of solvates(4). The case has been regarded as analogous to the distinction between the "red" and "brown" solutions of iodine in non-associated and associated solvents respectively, where solvation has also been invoked to explain the greater displacement of the absorption in the brown solutions(11, 12). In both bromine and iodine the absorption band is displaced between about 1000 and 2000  $\text{cm}^{-1}$  to higher frequencies; but in bromine the change in colour is not so obvious since the undisplaced band is already further to the violet. Solutions of bromine in ethanol at  $-70^\circ\text{C}$ . are pale yellow in colour, however (see the absorption curve in Fig. 1).

There are many arguments against solvation in these solutions, at least in so far as the spectroscopic evidence is concerned. In the first place, "solvation" is an ill-defined term; but if any real solvation in the chemical sense is meant, it must imply the formation of bonds that are strong compared with  $kT$ , say at least as strong as a hydrogen bond. If there is solvation of this kind in solutions of bromine in associated solvents, it is remarkable that there is so little difference between the spectra in associated and non-associated solvents, or between the former and the gas, in respect both of position and intensity. It is also difficult to understand why solvates in different associated solvents should have such similar spectra. If solvates are responsible for the displacement of 650  $\text{cm}^{-1}$  in sulphuric acid at room temperature, it is difficult to account for the further displacement of 500  $\text{cm}^{-1}$  when the acid is frozen to a glass or for the displacement of 1300  $\text{cm}^{-1}$  in pentane at low temperatures(5). If there is any kind of chemical equilibrium between solvated and unsolvated bromine molecules, one would expect that changes in concentration would cause opposite changes in the intensities of the band due to unsolvated molecules and that due to solvates with a constant value of  $\epsilon$  at the point where the absorption curves cross. There is no evidence of this behaviour in the spectra of solutions of either bromine or iodine.

Consideration of the data in Table 1 reveals a regular trend in the values of  $\epsilon_{max}$ ,  $\Delta\nu_{max}$ , and  $\Delta\nu_{\frac{1}{2}}$ . With the exception of irregularities in cyclohexane and carbon tetrachloride, and in sulphuric acid at  $-70^\circ\text{C}$ . which is discussed later, the half-width  $\Delta\nu_{\frac{1}{2}}$  increases regularly as  $\Delta\nu_{max}$  increases. There is also a parallel decrease in  $\epsilon_{max}$ , although in this case the non-associated and the

associated solvents must be regarded as two separate groups. This regularity of behaviour is more likely to result from the operation of some (probably imperfectly understood) physical factor than from a chemical solvation which would seem to imply some individuality in solvents of differing composition.

TABLE I  
DATA RELATING TO A MAXIMUM OF BROMINE IN SOLVENTS

Solvent	Temperature (° C.)	$\epsilon_{max}$	$\lambda_{max}$ (Å.)	$\nu_{max}$ (cm. <sup>-1</sup> )	Dis- place- ment $\Delta\nu_{max}$ (cm. <sup>-1</sup> )	Half- width $\Delta\nu_{\frac{1}{2}}$ (cm. <sup>-1</sup> )	Oscil- lator Strength (f)
(Gas) (a) .. ..	18	152	4150	24100	—	4400	0.0029
<i>n</i> -Hexane .. ..	18	226	4170	24000	-100	4200	0.0041
Cyclohexane (b) .. ..	18	216	4150	24100	0	5000	0.0047
Chloroform (b) .. ..	18	208	4100	24400	300	4600	0.0041
Carbon tetrachloride (b) ..	18	203	4065	24600	500	4500	0.0039
Sulphuric acid .. ..	18	212	4040	24750	650	5300	0.0048
Sulphuric acid .. ..	-70	238	3960	25250	1150	4700	0.0048
Phosphoric acid .. ..	18	181	4000	25000	900	5700	0.0044
Water (b) .. ..	18	164	3920	25500	1400	6300	0.0044
Ethanol .. ..	-70	(300) ?	3740 ?	26700 ?	2600 ?	?	?

(a) Data from Acton, Aickin, and Bayliss(6).

(b) Data from Aickin, Bayliss, and Rees(1).

Bayliss and Rees(13) have shown that the assumption of solvates in the chemical sense is unnecessary, and that the spectra of bromine and iodine solutions can be explained in terms of the cage of solvent molecules that encloses each solute molecule. Without assuming any special solvation forces but merely those physical forces that are responsible, for example, for the activation energy of diffusion, the presence of the cage introduces a potential energy barrier (of the order of magnitude of the activation energy of diffusion) which modifies the  $U(r)$  curves of both the normal and excited states concerned in the electronic transition. The corresponding wave functions are modified, and the application of the theory of continuous absorption spectra(14, 15) made it possible to account qualitatively for the displacements and the changes in symmetry of the absorption curves of bromine and iodine in non-associated solvents.

In associated solvents, the hydrogen bonding that is responsible for association leads to a certain preferred "hole" size in the solvent which is usually smaller than solute molecules such as bromine and iodine. In these cases, there must be mutual accommodation between solvent and solute which results in an expansion of the solvent "hole" and a slight compression of the solute molecule in its ground state. This slight compression is sufficient to account for the greater displacement of the absorption band since the Franck-Condon transition from the ground state is to a point of higher energy in the excited

state. Rees(16) has since obtained a quantitative evaluation of the work needed to expand the solvent cage in terms of the surface tension of the solvent, and has calculated displacements of the bands of iodine and bromine in water ( $1300\text{ cm.}^{-1}$  and  $1400\text{ cm.}^{-1}$  respectively) that are in good agreement with the experimental values ( $1800\text{ cm.}^{-1}$  and  $1400\text{ cm.}^{-1}$ \*).

The above interpretation of the effect of associated solvents leads to the qualitative predictions that the displacement of  $\nu_{max}$  should be less in associated solvents of greater natural hole size, and also that increasing the rigidity of the solvent (as in freezing) should increase the displacement. The latter effect is obvious in the absorption curves in sulphuric acid at room temperature and at  $-70^\circ\text{C}$ . (Fig. 1). It can be shown that the former effect is realized in solutions in sulphuric and phosphoric acids as compared with water, although the data on these solvents are not sufficient to carry out a quantitative calculation.

The normal hole size in water is about  $3.4\text{ \AA}$ . in diameter(17), which must be expanded to accommodate a bromine molecule of diameter about  $5\text{ \AA}$ . Concentrated sulphuric and phosphoric acids have been shown by X-ray methods(18) to contain tetrahedral  $\text{SO}_4^{--}$  and  $\text{PO}_4^{---}$  ions associated by hydrogen bonds. The distance between the hydrogen bonded O atoms is  $2.85\text{ \AA}$ . and the angle between the H-O and the S-O or P-O bonds is given as tetrahedral. Using Pauling's(19) values of  $1.51\text{ \AA}$ . and  $1.55\text{ \AA}$ . for the S-O and P-O distances respectively in the ions, it can be deduced that the closest approach of two S nuclei in sulphuric acid is  $3.86\text{ \AA}$ ., and of two P nuclei in phosphoric acid is  $3.90\text{ \AA}$ . These values lead to effective packing radii of  $1.93\text{ \AA}$ . and  $1.95\text{ \AA}$ . for the  $\text{SO}_4^{--}$  and  $\text{PO}_4^{---}$  ions respectively. The effective packing radius of the  $\text{H}_2\text{O}$  molecule in water is  $1.38\text{ \AA}$ .(17). Hence one concludes that the normal hole size in the two acids is considerably greater than in water. Their surface tensions ( $63\text{ dyne cm.}^{-1}$  for 90 per cent. sulphuric acid) are not very different from that of water ( $73\text{ dyne cm.}^{-1}$ ). It is therefore to be expected that the compression of the bromine molecule will be less in the acid solvents than in water, and consequently the displacement of the absorption maximum will also be less, as observed.

### *(b) The Effect of Solvents on the Intensity*

Dealing first with the particular case of the solutions in sulphuric acid, it is to be noted that although  $\epsilon_{max}$  is increased from 212 to 238 on going from room temperature to  $-70^\circ\text{C}$ ., the total intensity of the band as measured by its oscillator strength  $f$  is the same at the two temperatures. Acton, Aickin, and Bayliss(6) have shown that the absorption spectrum of bromine gas is the sum of contributions from molecules distributed in statistical equilibrium between the various vibrational states  $v''=0$ ,  $v''=1$  etc. These contributions obey a sum rule so that the total intensity of the electronic transition is independent of temperature, even though the value of  $\epsilon_{max}$  and the breadth of the band may change.

\* Rees quotes  $1750\text{ cm.}^{-1}$  for this displacement from Aickin, Bayliss, and Rees(1). A re-examination of the data of these authors gives  $1400\text{ cm.}^{-1}$  as a better value.



At low temperatures, the contribution from the state  $v''=0$  becomes more important, resulting in greater  $\epsilon_{max}$ , a narrower absorption band, and a more pronounced hump on the low frequency side due to the  $B$  component of the absorption (cf. absorption curve of the state  $v''=0$ , Acton, Aickin, and Bayliss,6). These features are evident in the comparison of the absorption curves in Figure 1 for sulphuric acid at the two temperatures.

A statistical calculation at  $-70^\circ\text{C}$ . shows that practically all the molecules in bromine would be in the vibrational states  $v''=0$  (89.6 per cent.) and  $v''=1$  (10.4 per cent.). Using the extinction coefficients determined by Acton, Aickin, and Bayliss(6) for these states, it is calculated that  $\epsilon_{max}$  in bromine gas at  $-70^\circ\text{C}$ . would be increased 10 per cent. over its value at room temperature. Table 1 shows that in sulphuric acid,  $\epsilon_{max}$  is increased 12 per cent. at  $-70^\circ\text{C}$ . It is clear from these results and from the constant value of  $f$  that the temperature effect in sulphuric acid solution is explained by changes in the statistical distribution of bromine molecules between their various vibrational states, whatever may be the reason for the overall increase in intensity of nearly 70 per cent. as compared with the gas.

Chako(20) has considered the relationship between the  $f$  values of absorption bands in the gas state and in solution on the basis of classical dispersion theory, and in terms of the clarification of his work by Mulliken and Rieke(10) has shown that the existence of the Lorentz-Lorenz forces in solution requires the factor  $\gamma$  in equation (1) to have the value  $9n/(n^2+2)^2$ , if  $n$  is the refractive index of the solvent. The solvents in Table 1 have refractive indices lying roughly between 1.4 and 1.5;  $\gamma$  is therefore between 0.8 and 0.75, and the resulting values of  $f$  are between 0.8 and 0.75 of the values quoted in Table 1. Chako's theory is thus able to account for about half of the observed difference in intensity between absorption in the gas state and in solution. The remainder of the difference cannot be explained precisely in view of the present state of the theory of absorption spectra in solution. It may be due to the inaccuracy of the classical dispersion theory used by Chako, or else to solvent perturbations modifying the selection rules of the weak or forbidden transitions that are responsible for the spectrum(10).

## V. ACKNOWLEDGMENTS

We wish to acknowledge support from the Commonwealth Research Grant to Australian Universities in the form of a studentship (A.R.H.C.) and from the Hackett Fund of the University of Western Australia in the form of a research scholarship (B.G.G.).

## VI. REFERENCES

- (1) AICKIN, R. G., BAYLISS, N. S., and REES, A. L. G.—*Proc. Roy. Soc. A* **169**: 234 (1938).
- (2) BOVIS, P.—*Ann. Phys. Paris* **10**: 232 (1928).
- (3) GILLAM, A. E., and MORTON, R. A.—*Proc. Roy. Soc. A* **124**: 604 (1929).



- (4) CHILD, C. L., and WALKER, O. J.—*Trans. Faraday Soc.* **34**: 1506 (1938).
- (5) PRIKHOTKO, A.—*Acta Physicochim. U.R.S.S.* **16**: 125 (1942).
- (6) ACTON, A. P., AICKIN, R. G., and BAYLISS, N. S.—*J. Chem. Phys.* **4**: 474 (1936).
- (7) BAYLISS, N. S., and REES, A. L. G.—*Ibid.* **7**: 854 (1939).
- (8) REES, A. L. G.—*Proc. Phys. Soc.* **59**: 1008 (1947).
- (9) MULLIKEN, R. S.—*Phys. Rev.* **57**: 500 (1940).
- (10) MULLIKEN, R. S., and RIEKE, C. A.—*Rep. Progr. Phys.* **8**: 231 (1941).
- (11) WALKER, O. J.—*Trans. Faraday Soc.* **31**: 1432 (1935).
- (12) FAIRBROTHER, F.—*Nature*, **160**: 87 (1947).
- (13) BAYLISS, N. S., and REES, A. L. G.—*J. Chem. Phys.* **8**: 377 (1940).
- (14) GIBSON, G. E., RICE, O. K., and BAYLISS, N. S.—*Phys. Rev.* **44**: 193 (1933).
- (15) BAYLISS, N. S.—*Proc. Roy. Soc. A* **158**: 551 (1937).
- (16) REES, A. L. G.—*J. Chem. Phys.* **8**: 429 (1940).
- (17) BERNAL, J. D., and FOWLER, R.—*Ibid.* **1**: 515 (1933).
- (18) FINBAK, C., RØNNING, O., and VIERVOLL, H.—*Tidsskr. Kjemi Bergv. Met.* **4**: 26 (1944).  
(*Chem. Abstr.* **40**: 3034 (1946).)
- (19) PAULING, L.—“The Nature of the Chemical Bond.” (Cornell Univ. Press: Cornell, 1940.)
- (20) CHAKO, N. Q.—*J. Chem. Phys.* **2**: 644 (1934).

# THE FREE ENERGIES OF HYDRATION OF GASEOUS IONS

By N. S. HUSH\*

[Manuscript received August 27, 1948]

## Summary

Values of hydration energies of individual ions have usually been obtained by division of sums of energies of hydration of pairs of ions, and those calculated by different authors are usually mutually inconsistent. "Experimental" figures, whenever these are quoted, have always been obtained by assuming the truth of theoretical equations whose accuracy has not been independently checked. The distinction between free energy of ion/water-molecule interaction and the real free energy of hydration of a gaseous ion is pointed out, and the importance of Klein and Lange's measurement of the Volta-potential  $\text{Hg}/\text{Hg}^+$  (soln.), which makes possible the direct calculation of real free energies of hydration of individual ions, thus providing a check on theoretical values, is emphasized.

Utilizing this value, the equation  $-\Delta F_h^\circ = -\Delta F_f^\circ + \Delta F_i^\circ + \Delta F_s^\circ - 103.92z \text{ kcal.}$  (where  $\Delta F_s^\circ$  is the free energy of formation of the gaseous monatomic element,  $\Delta F_i^\circ$  is the free energy of ionization,  $\Delta F_f^\circ$  is the free energy of formation of the aqueous ion, and  $\Delta F_h^\circ$  is the real free energy of hydration of the ion, of valency  $z$ , at  $298.2^\circ \text{ K.}$ ) is derived from fundamental considerations. By means of this equation, the real free energies of hydration of 49 ions are calculated, using the most reliable data. It is proposed that these be provisionally accepted as standard values. Several subsidiary values for important ions are calculated indirectly. The difference between  $\Delta F_h^\circ$  and the free energy of ion/water-molecule interaction is discussed in relation to the surface structure of water: a value of  $-0.30 \text{ v.}$  is derived for the  $\chi$ -potential at the surface of pure water, and it is concluded that at the water/gas interface the positive poles of the surface layer are oriented towards the gas phase. The applicability of a modified Born equation in the calculation of free energies of hydration is discussed, and a modified equation is proposed which yields values of  $\Delta F_h^\circ$  for gaseous ions with noble gas structure in excellent agreement with those calculated independently by the method described above.

## I. INTRODUCTION

During the last two decades, several discussions of the mechanism of hydration of gaseous ions and calculations of hydration energies of individual ions have been published.† Almost without exception, the more recent calculations have proceeded either by considering the mutual potential energies of ions and water molecules on the basis of a particular theory of the charge distribution within the water molecule, or by employing a modified Born equation,

$$-\Delta F = \frac{Nz^2e^2}{2r_e}(1-1/D),$$

\* Department of Organic and Applied Chemistry, University of Sydney.

† The more important papers are listed in references (1-13).

assuming that water behaves as if it were a normal continuous dielectric (with dielectric constant approaching the macroscopic value) up to the surface of a spherical cavity containing the ion, whose radius  $r_e$  is simply related to the crystal radius of the ion. In each case, the quantity measured for an individual ion is the energy of interaction of ion and water molecules, the "chemical" free energy of hydration in the terminology of Lange(14). In each case, also, the "chemical" free energies of hydration of individual ions are obtained by a division of sums of energies of hydration of pairs of ions. From the thermal data (lattice energy and heat of solution) from which such sums are derived, it is not possible directly to measure the hydration energy of an individual ion, although this fact is sometimes obscured by comparison of "observed" with "calculated" values of ionic hydration energies. Thus Bernal and Fowler(6) provide a set of "observed" values of heats of hydration, to be compared with their calculated values, but these values are calculated on the basis of precisely the same assumptions regarding the mutual potential energies of ions and water molecules as were made in deriving the "calculated" values—assumptions which have in fact been conclusively shown by Verwey(12, 13) to be false and to lead to large errors. (The error in anions approaches 30 per cent., if Verwey's own figures are correct.) Again, Voet(7) argues that the "empirical relation" between ionic radii and heats of hydration of anions does not accord with the modified Born-Bjerrum equation, and, because of this, questions the applicability of the equation; but in fact all that he has shown is that the theory underlying the calculations of van Arkel and de Boer(5) (whose values for ionic heats of hydration Voet refers to) is incompatible with that underlying the application of the Born equation: the findings of van Arkel and de Boer are not themselves criticized, and thus no warrant is given for referring to them, rather than those obtained by applying the Born equation, as "empirical".

The position is different with real free energies of hydration. The real free energy of hydration of an ion may be defined as the real free energy of removal of an ion from a point *in vacuo* outside the action of surface and induction forces to the interior of the aqueous phase: this (unlike the "chemical" hydration energy) includes the work done against the surface potential of the phase when the ion crosses the boundary. It has been shown by Klein and Lange(8, 9) that real free energies of hydration of individual ions can be determined directly—i.e. without reference to lattice energies or heats of solutions of salts—when the electrochemical normal potential  $M/M^+$  on the conventional hydrogen scale and the Volta-potential  $M/M^+$  (soln.) are known. Klein and Lange derived values of the real free energy of hydration of several ions by way of intermediary calculation of the separate Volta-potentials. This procedure is, however, unduly complicated. In what follows, a simple and general relationship between free energy of hydration of the gaseous ion and the free energies of sublimation and ionization of the element, and the free energy of formation of the aqueous ion, is derived, employing Klein and Lange's value of the Volta-potential  $Hg/Hg^+$  (soln.). This has been employed to calculate the free energies of hydration of 49 common ions, using the most reliable data. It is proposed that these be provisionally accepted as standard values.

The figures obtained differ from those of Klein and Lange partly because of the use of different data, and partly because of insufficient distinction between heat and free energy terms in these authors' calculations. In one important case, there are different grounds for disagreement: Klein and Lange have calculated a value for the energy of hydration of  $\text{Hg}^+(\text{g})$ , employing the measured value of the mercury/mercurous potential. This is erroneous, for the mercurous ion is completely dimerized at all measurable concentrations in aqueous solution, and since the free energy of dimerization of the ion is not known, the energy of hydration cannot be calculated. An incorrect calculation of the energy of hydration of  $\text{Hg}^+(\text{g})$  has also been made by Cartledge(24).

The direct calculation of chemical free energies of hydration will only be possible when accurate values of solution surface potentials are known. However, an indirect calculation from real free energies of hydration is possible, as is shown below. It may be noted that in many calculations in which the free energy of hydration is necessary (particularly those of electrochemical potentials from thermal data) the real, rather than the chemical, free energies of hydration must be known.

## II. THEORETICAL DERIVATION

### (a) General

We shall consider an electrochemical system of the type :

1	2	3	4	1'
$\text{M}_1$	$\text{M}_1^+, \text{H}_2\text{O}$ $A^-$	$\text{M}_2^+, \text{H}_2\text{O}$ $A^-$	$\text{M}_2$	$\text{M}_1$

(The separate phases are designated by numerals.)

Where liquid-junction potentials are eliminated, the e.m.f. of the cell is related to the free energies\* of hydration ( $\Delta F_h^\circ$ ), sublimation ( $\Delta F_s^\circ$ ), and ionization ( $\Delta F_i^\circ$ ) of  $\text{M}_1$  and  $\text{M}_2$  by

$$\text{e.m.f.} = \Delta F_s^\circ_{(1)} + \Delta F_i^\circ_{(1)} - \Delta F_h^\circ_{(1)} - (\Delta F_s^\circ_{(2)} + \Delta F_i^\circ_{(2)} - \Delta F_h^\circ_{(2)}) \quad (1)$$

Expressing the e.m.f. as difference of the normal potentials  $E_{0(1)}$  and  $E_{0(2)}$  of the couples  $\text{M}_1/\text{M}_1^+$  and  $\text{M}_2/\text{M}_2^+$  we may write

$$E_{0(1)} = \Delta F_s^\circ_{(1)} + \Delta F_i^\circ_{(1)} - \Delta F_h^\circ_{(1)} - (\Delta F_s^\circ_{(2)} + \Delta F_i^\circ_{(2)} - \Delta F_h^\circ_{(2)}) - E_{0(2)} \quad (2)$$

If for any metal  $\text{M}_2$ , the values of  $\Delta F_s^\circ$ ,  $\Delta F_i^\circ$ ,  $\Delta F_h^\circ$ , and  $E_0$  were known, we could thus write† for any other metal  $\text{M}_1$

$$-\Delta F_h^\circ = -\Delta F_f^\circ + \Delta F_s^\circ + \Delta F_i^\circ + C \quad (3)$$

\* For brevity, it is assumed (unless otherwise stated) that the free energy terms are expressed in electron-volts.

† An equation of approximately the same form as equation (3) was employed by Webb(4) to calculate subsidiary values of free energies of hydration.



where  $\Delta F_f^\circ$  is the free energy of formation of the aqueous  $M_1^+$  ion and  $C$  is a constant, given by

$$C = -\Delta F_s^\circ(2) - \Delta F_i^\circ(2) - \Delta F_h^\circ(2) + \Delta F_f^\circ(2) \quad \dots\dots\dots (4)$$

where  $\Delta F_f^\circ(2)$  is the free energy of formation of aqueous  $M_2^+$ .

### (b) Evaluation of $C$

By expressing the e.m.f. of the above system as the sum of the appropriate Volta-potentials\*, we may obtain a value for  $C$ .

The e.m.f. (assuming the liquid-junction potential to be negligible, as before) may be expressed as the sum of three Volta-potentials

$$\text{e.m.f.} = \Delta\psi_{(12)} + \Delta\psi_{(34)} + \Delta\psi_{(41')} \quad \dots\dots\dots (5)$$

where  $\Delta\psi_{(12)}$  is the Volta-potential  $M_1/M_1^+$  (soln.),  $\Delta\psi_{(34)}$  is the Volta-potential  $M_2^+$  (soln.)/ $M_2$ , and  $\Delta\psi_{(41')}$  is the Volta-potential  $M_2/M_1$ . (All values refer to unit activity.) We shall choose Hg as the reference metal  $M_2$ , since the potential  $\Delta\psi_{(43)}$  Hg/Hg<sup>+</sup> (soln.) has been measured by Klein and Lange.

$\Delta\psi_{(12)}$  may be equated to  $(-\Delta F_h^\circ(1) + \text{free energy of removal of a metal ion from } M_1 \text{ in vacuo})$ . The free energy of removal of  $M_1^+$  (g) from  $M_1$  (crystall.) is clearly given by  $(\Delta F_s^\circ - \Delta F_i^\circ - \varphi_1)$ , where  $\varphi_1$  is the work of electron removal of  $M_1$ . Thus

$$\Delta\psi_{(12)} = -\Delta F_h^\circ(1) + \Delta F_s^\circ + \Delta F_i^\circ - \varphi_1 \quad \dots\dots\dots (6)$$

The Volta-potential  $\Delta\psi_{(41')}$  is simply the difference of work of electron removal  $\varphi_1$  and  $\varphi_2$  of  $M_1$  and  $M_2$ , i.e.

$$\Delta\psi_{(41')} = -\varphi_2 + \varphi_1 \quad \dots\dots\dots (7)$$

Substituting in equation (5), we have

$$\text{e.m.f.} = -\Delta F_h^\circ(1) + \Delta F_s^\circ(1) + \Delta F_i^\circ(1) + \Delta\psi_{(34)} - \varphi_2 \quad \dots\dots\dots (8)$$

Substituting the free energies of formation of the aqueous ions for the electrochemical potentials, we obtain

$$-\Delta F_h^\circ(1) = -\Delta F_f^\circ(1) + \Delta F_s^\circ(1) + \Delta F_i^\circ(1) + (\Delta\psi_{(34)} - \varphi_2 + \Delta F_f^\circ(2)) \quad \dots\dots\dots (9)$$

From equations (3), (4), and (9) it is evident that

$$C = \Delta F_s^\circ(2) - \Delta F_i^\circ(2) - \Delta F_h^\circ(2) + \Delta F_f^\circ(2) = \Delta\psi_{(34)} - \varphi_2 + \Delta F_f^\circ(2) \quad \dots\dots\dots (10)$$

In order to obtain the value of  $C$ , we shall take the values  $\Delta\psi_{(34)}$ ,  $\varphi_2$ , and  $\Delta F_f^\circ(2)$  to be  $-0.69$  v.†,  $-4.52$  v.†, and  $0.80$  v.‡ respectively. Substitution of these in equation (10) yields the important value

$$C = -4.41 \text{ v.} \quad \dots\dots\dots (11)$$

Substitution in equation (9) yields the equations (valid for any univalent cation)

$$-\Delta F_h^\circ = -\Delta F_f^\circ + \Delta F_i^\circ + \Delta F_s^\circ - 4.41 \quad \dots\dots\dots (12)$$

\* The theory of Volta (phase-boundary) potentials is discussed by Lange(14), Frumkin(15), Klein and Lange(8, 9), and Chalmers(16), and has been reviewed by Burgers(17).

† Klein and Lange(8).

‡ Latimer(10).

where the free energy terms are expressed in electron-volts, and

$$-\Delta F_h^\circ = -\Delta F_f^\circ + \Delta F_i^\circ + \Delta F_s^\circ - 103.92 \dots\dots (13)$$

where these are expressed in kilocalories.

(c) *Generalization of Equation (13)*

A more general equation, holding both for cations and for anions of any valency, may be obtained by the substitution of  $zC$  for  $C$  in equation (13), where  $z$  is the valency of the ion (e.g. for  $S^{=}$ ,  $z = -2$ ). This yields as the final expression

$$-\Delta F_h^\circ = -\Delta F_f^\circ + \Delta F_i^\circ + \Delta F_s^\circ - 103.92z \dots\dots\dots (14)$$

where the free energy terms are expressed in kilocalories. In multivalent cations,  $\Delta F_i^\circ$  is derived from the sum of the first  $n$  ionization potentials of the ion; and for anions,  $\Delta F_i^\circ$  is derived from the corresponding electron affinity.

### III. TABULATION OF DATA

The free energies of hydration of 49 ions are listed in Table 1. All have been calculated by means of equation (14) and the values of  $\Delta F_f^\circ$ ,  $\Delta F_s^\circ$ , and  $\Delta F_i^\circ$  used in each calculation are shown; the sources of these values are indicated. The Goldschmidt radius of the ion(18) (where it is known) is also listed.

(a) *Free Energies of Formation*

Where not otherwise indicated, values of  $\Delta F_f^\circ$  have been taken from the compilation of Latimer(19), or calculated from the  $\Delta H_f^\circ$  values of Bichowsky and Rossini(20). In the latter case the entropy values of Latimer *et al.* listed in (19) have been assumed.

(b) *Free Energies of Ionization*

These were obtained from values of ionization potentials (employing the figures of Sherman, 31), and electron affinities of positive and negative ions respectively. In calculating  $\Delta F_i^\circ$  (298.2) it is assumed (i) that within experimental limits these figures give  $\Delta H_0^\circ$ , (ii) that the three species in the equilibrium  $M(g) \rightleftharpoons Mg^{z+} + ze(g)$  are perfect monatomic gases.\* Then for a single ionization  $\Delta C_p \approx 5$  and the general equation is

$$\Delta F_i^\circ (298.2) = \Delta H_0^\circ - 5T \ln T + IT.$$

Employing the equation of Lewis and Randall(22) to calculate  $\Delta S^\circ$ , viz.

$$\Delta S^\circ (t) = \Delta \Gamma_0(1 + \ln T) + \Delta \Gamma_1 T + \frac{1}{2} \Delta \Gamma_2 T^2 \dots - I$$

we obtain

$$\Delta S^\circ (t) = 5(1 + \ln T) - I.$$

The value of  $\Delta S^\circ$  was calculated by assuming the values of  $S^\circ$  (298.2) of gaseous ion and electron gas to be given by the Sackur equation

$$(S^\circ (298.2) = 3/2 R \ln W + 26.03 \text{ cal./deg.})$$

and the entropy of the gaseous element to be given by this equation plus a multiplicity term ( $R \ln M$ ). This yields

$$\Delta S^\circ (298.2) = 3.61 - 1.33 = 2.28 \text{ cal./deg.}$$

\* This treatment differs somewhat from that of Lewis and Randall(22).

This yields the value 31.20 for the constant  $I$ . Thus

$$\begin{aligned}\Delta F_{(298.2)}^\circ &= \Delta H_0^\circ - 5T \ln T + 31.20T \\ &= \Delta H_0^\circ + 0.80 \text{ kcal.}\end{aligned}$$

By similar reasoning\* the values of  $(\Delta F_{(298.2)}^\circ - \Delta H_0^\circ)$  for divalent, trivalent, and tetravalent ions are found to be 1.46, 2.04, and 2.35 kcal. respectively.

### (c) Free Energies of Sublimation

Values of  $\Delta F_s^\circ$  have been taken, unless otherwise indicated, from Latimer(19) or calculated from the values of  $\Delta H_s^\circ$  of Bichowsky and Rossini(20). The entropy values are in general those of Latimer(19) but in some cases the more recent values in the compilation of Kireev(23) have been used.

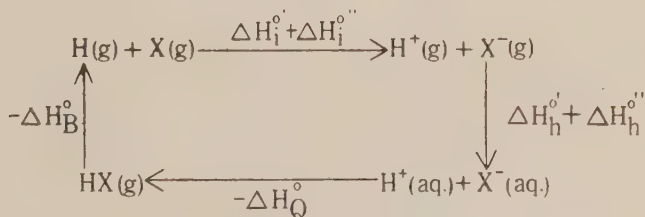
*Accuracy.*—The limit of accuracy in the determination of the Volta-potential  $\text{Hg}/\text{Hg}^+$  (soln.) has been estimated at  $\pm 0.03$  v.; this will therefore contribute  $\pm 0.6$  kcal. to the possible error of the values of  $\Delta F_h^\circ$  listed in Table 1. The limits of accuracy of other quantities will in general be larger than this, the least accurate of all the data being those of the free energies of sublimation or dissociation. For the alkali metal and halogen ions, however, the possible error may be estimated at  $\pm 2.0$  kcal.†

## IV. SUBSIDIARY VALUES

Employing the data of Table 1 together with other thermal data, it is possible to make alternative estimates of free energies of hydration: as an example, the free energy of hydration of the proton is recalculated below. Similar methods are used to calculate the free energies of hydration of  $\text{NH}_4^+(\text{g})$ ,  $\text{OH}^-(\text{g})$  and  $\text{SH}^-(\text{g})$ .

### (a) Free Energy of Hydration of the Proton

The heat of hydration of the proton may be independently calculated from thermal data for the hydrogen halides, by means of the following cycle:



\* The multiplicity data are those of Herzberg(21).

† The experimental difficulties in measuring 4-potentials of this sort are considerable and this accuracy should perhaps not be insisted upon. As the maximum accuracy claimed for the hydration energies calculated in Table 1 is (for the alkali metal and halogen ions) approximately 3.5 times greater than the estimated error in the 4-potentials, the final result would not be much affected by a somewhat lower limit of accuracy in the measurement of the latter quantities.

where  $\Delta H_B^\circ$  is the bond strength,  $\Delta H_i^{\circ'}$  and  $\Delta H_i^{\circ''}$  are the heats of ionization of H(g) and X-(g) respectively,  $\Delta H_h^{\circ'}$  and  $\Delta H_h^{\circ''}$  are the heats of hydration of H<sup>+</sup>(g) and X-(g) respectively, and  $\Delta H_Q^\circ$  is the heat of solution of HX(g). Such a cycle was used by van Arkel and de Boer(5) to calculate  $\Delta H_h^{\circ'}$  from the data for HCl, and by Baughan(35) to calculate this value from data for each of

TABLE I  
REAL FREE ENERGIES OF HYDRATION OF GASEOUS IONS AT 298.2° K.

Ion	$r_a$ (Å)	$-\Delta F_f^\circ$ (kcal.)	$\Delta F_s^\circ$ (kcal.)	$\Delta F_i^{\circ a}$ (kcal.)	$-\Delta F_h^\circ$ (kcal.)
<i>Univalent—</i>					
Cs <sup>+</sup> ..	1.65	70.28	12.23	90.23	68.8
Rb <sup>+</sup> ..	1.49	68.80	13.35	96.65	74.9
K <sup>+</sup> ..	1.33	67.43	14.64	100.47	78.6
Na <sup>+</sup> ..	0.98	62.59	18.64	118.81	96.1
Li <sup>+</sup> ..	0.78	70.70	28.81	124.46	120.1
Au <sup>+</sup> ..	1.37	-38.70	82.50	212.78	152.7
Ag <sup>+</sup> ..	1.13	-18.44	58.70	174.49	110.8
Cu <sup>+</sup> ..	0.96	-12.04	71.70	178.18	108.3
Tl <sup>+</sup> ..	1.49	7.76	31.70	141.04	76.6
H <sup>+</sup> ..	—	0.00	48.35	312.88	257.3
I <sup>-</sup> ..	2.20	12.33	16.87	-70.32 <sup>a</sup>	62.8
Br <sup>-</sup> ..	1.96	24.58	19.90	-77.60 <sup>a</sup>	70.8
Cl <sup>-</sup> ..	1.81	31.33	25.09	-83.14 <sup>a</sup>	77.2
F <sup>-</sup> ..	1.33	65.70	27.60	-90.52 <sup>a</sup>	106.7
<i>Divalent—</i>					
Be <sup>2+</sup> ..	0.34	78.70	66.00	633.93	570.8
Mg <sup>2+</sup> ..	0.78	108.76 <sup>b</sup>	27.64	522.29	450.9
Ca <sup>2+</sup> ..	1.06	132.70	34.51	414.57	374.0
Sr <sup>2+</sup> ..	1.27	133.20	30.70	385.51	341.6
Ba <sup>2+</sup> ..	1.43	133.85	41.40	350.68	318.1
V <sup>2+</sup> ..	—	70.00	75.00	496.46	433.7
Cr <sup>2+</sup> ..	—	39.40	78.63	539.59	449.8
Mn <sup>2+</sup> ..	0.91	50.60	59.54	533.99	436.3
Fe <sup>2+</sup> ..	0.83	20.31	85.68	562.66	460.9
Co <sup>2+</sup> ..	0.82	12.80	74.00	580.65	459.7
Pd <sup>2+</sup> ..	—	-45.70 <sup>c</sup>	101.00	649.62	497.1
Ni <sup>2+</sup> ..	0.78	11.53	87.43	595.09	486.3
Ag <sup>2+</sup> ..	—	-63.15	58.70	569.58	357.3
Cu <sup>2+</sup> ..	0.80 <sup>d</sup>	-15.91	71.70	644.77	492.8
Hg <sup>2+</sup> ..	1.12	-39.42	7.59	671.76	432.1
Cd <sup>2+</sup> ..	1.03	18.55	18.49	596.33	425.6
Zn <sup>2+</sup> ..	0.83	35.18	22.69	629.78	479.9
Pb <sup>2+</sup> ..	1.32	5.81	39.60	517.22	354.8
Sn <sup>2+</sup> ..	1.04 <sup>e</sup>	6.28	69.70	506.38	374.6
Se <sup>2-</sup> ..	1.91	-35.76	51.40 <sup>f</sup>	97.0	320.5
S <sup>2-</sup> ..	1.74	-23.42	56.60 <sup>f</sup>	79.0	320.0



TABLE 1—Continued

Ion	$r_a$ (Å)	$-\Delta F_f^\circ$ (kcal.)	$F\Delta_s^\circ$ (kcal.)	$\Delta F_i^{\circ a}$ (kcal.)	$-\Delta F_h^\circ$ (kcal.)
<i>Trivalent—</i>					
La <sup>3+</sup> ..	1.22	163.80	80.00	864.71	796.7
Y <sup>3+</sup> ..	1.06	145.00	80.30	910.84	824.3
Sc <sup>3+</sup> ..	0.83	138.00	60.00	1017.14	843.4
Al <sup>3+</sup> ..	0.57	115.50	57.79	1225.23	1086.7
V <sup>3+</sup> ..	0.65	75.90 <sup>b</sup>	75.00	1091.04	860.1
Cr <sup>3+</sup> ..	0.64	49.00	78.63	1174.0 <sup>h</sup>	989.9
Fe <sup>3+</sup> ..	0.67	2.53	85.68	1350.0 <sup>h</sup>	1124.5
Co <sup>3+</sup> ..	—	-28.90	74.00	1390.0 <sup>h</sup>	1123.3
Ti <sup>3+</sup> ..	1.05	-49.74	31.70	1295.58	965.7
In <sup>3+</sup> ..	0.92	23.50	44.00	1212.55	968.3
Ga <sup>3+</sup> ..	0.62	36.00	43.20	1281.05	1048.5
<i>Tetravalent—</i>					
Th <sup>4+</sup> ..	1.10	-72.20	39.60	2267.0 <sup>h</sup>	1819.2
Pb <sup>4+</sup> ..	0.84	197.00	167.50	1602.35	1499.2
Sn <sup>4+</sup> ..	0.74	-0.65	69.70	2146.34	1799.8

<sup>a</sup> The values of  $\Delta F_i^\circ$  for the halogens were obtained indirectly. The heats of hydration of pairs of ions  $M^+$  and  $X^-$  (where  $M^+ = \text{Li}^+, \text{Na}^+, \text{K}^+, \text{Rb}^+, \text{Cs}^+$ ) were calculated from the crystal energies of Huggins(25), corrected to 298.2° K., and the heat of solution of the corresponding salt. Several values of  $-\Delta H_h^\circ$  for each halogen ion were then obtained by subtracting the heat of hydration of the cation (calculated from the  $-\Delta F_h^\circ$  values of Table 1) and the mean taken. From the means,  $-\Delta F_h^\circ$  was calculated for each halogen ion, and the value of  $\Delta F_i^\circ$  corresponding to this figure determined. The entropies of hydration of both alkali metal and halogen ions have been assumed to be those given by Latimer, Pitzer, and Slansky(11).

<sup>b</sup> Stephenson(27).

<sup>c</sup> Templeton, Watt, and Garner(28).

<sup>d</sup> Approximate estimate by Goldschmidt(29).

<sup>e</sup> Kapustinskii(30).

<sup>f</sup> Sherman(31).

<sup>g</sup> Calculated from the  $V^{2+}/V^{3+}$  potential measurements of Jones and Colvin(32).

<sup>h</sup> Estimated values.

the hydrogen halides. In view of the importance of this figure, the calculation is repeated below: the data of Table 1, together with the values of Latimer, Pitzer, and Slansky(11) for the entropy of hydration of the halide ions,  $\Delta S_h^{\circ\prime\prime}$ , are used to calculate  $\Delta H_h^{\circ\prime\prime}$ . This is shown in Table 2.

The mean of the values of  $-\Delta H_h^{\circ\prime}$  shown in Table 2 is 267.3 kcal./g. ion.\* Taking the entropy of hydration of the proton as 26.0 cal./deg., the mean free

\* It is interesting to note that the earliest estimate, viz. that of Fajans(1), differs from this by less than 2 per cent., while the most recent estimate by Baughan(35),  $282.5 \pm 3$  kcal., is considerably in error: Webb calculated 249.6 kcal., and Garriek 255 kcal.

TABLE 2

INDEPENDENT CALCULATION OF THE HEAT OF HYDRATION OF THE PROTON

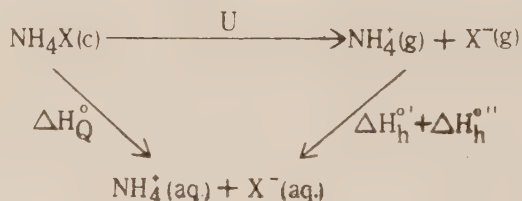
Halide	$\Delta H_B^{\circ*}$ (kcal.)	$\Delta H_i^{\circ'}$ (kcal.)	$\Delta H_i^{\circ''}$ (kcal.)	$\Delta S_h^{\circ''}$ (cal./deg.)	$-\Delta H_h^{\circ''}$ (kcal.)	$\Delta H_Q^{\circ}$ (kcal.)	$-\Delta H_h^{\circ'}$ (kcal.)
HF ...	147.5	-312.1	89.7	29	115.6	11.7	266.0
HCl ..	102.7	-312.1	82.3	15	81.7	17.6	268.3
HBr ..	87.3	-312.1	76.8	12	74.4	20.0	268.2
HI ..	71.4	-312.1	69.5	7	64.9	19.3	266.7

\* The bond energy data are those of Pauling(33).

energy of hydration of the proton is thus 259.5 kcal./g. ion. This compares within 1 per cent. with the value calculated in Table 1 by means of equation (14) (257.3 kcal./g. ion).

(b) *The Free Energy of Hydration of the Ammonium Ion*

A good estimate of the heat of hydration of  $\text{NH}_4^+$  may be obtained by means of the familiar cycle :



The calculation for the four halides is shown in Table 3.

TABLE 3

HEAT OF HYDRATION OF THE AMMONIUM ION AT 298.2° K.

Salt	$U_{298.2}^*$ (kcal.)	$\Delta H_Q^{\circ}$ (kcal.)	$\Delta H_h^{\circ''}$ (kcal.)	$\Delta H_h^{\circ'}$ (kcal.)
$\text{NH}_4\text{F}$ ..	177.5	1.5	115.6	60.4
$\text{NH}_4\text{Cl}$ ..	153.3	3.9	81.7	67.7
$\text{NH}_4\text{Br}$ ..	147.4	4.5	74.2	63.7
$\text{NH}_4\text{I}$ ..	143.6	3.6	64.9	75.1

\* Values of  $U$  are those theoretically calculated by Sherman(31).

The mean of these rather discordant values for  $\Delta H_h^\circ$  is  $-66$  kcal./g. ion. Taking the entropy of hydration to be  $8.2$  cal./deg., this yields  $-64$  kcal./g. ion for the free energy of hydration of  $\text{HN}_4^+$ .

(c) *The Free Energy of Hydration of the Hydroxyl Ion*

The free energy of hydration of  $\text{OH}^-$  may be obtained also by a simple cycle: the separate reactions, together with values of  $\Delta F^\circ$ , are shown below.

	$\Delta F^\circ_{298.2}$ (kcal.)			
$\text{H(g)} + \text{OH(g)} \longrightarrow \text{H}_2\text{O(g)}$	..	..	..	$-102.4$
$\text{H}^+(\text{g}) + e \longrightarrow \text{H(g)}$	..	..	..	$-312.1$
$\text{OH}^-(\text{g}) \longrightarrow \text{OH(g)} + e$	..	..	..	$59.0$
$\text{H}^+(\text{aq.}) \longrightarrow \text{H(g)}$	..	..	..	$257.3$
$\text{H}_2\text{O(l)} \longrightarrow \text{H}^+(\text{aq.}) + \text{OH}^-(\text{aq.})$	..	..	..	$19.5$
$\text{H}_2\text{O(g)} \longrightarrow \text{H}_2\text{O(l)}$	..	..	..	$-2.1$
<hr/>				
$\text{OH}^-(\text{g}) \longrightarrow \text{OH}^-(\text{aq.})$	..	..	..	$-81$

The value thus calculated for the free energy of hydration of  $\text{OH}^-$  is  $-81$  kcal./g. ion.

(d) *Free Energy of Hydration of the Hydrosulphide Ion*

The free energy of hydration of  $\text{HS}^-$  may be obtained by a series of steps analogous to those used in determining the value for  $\text{OH}^-$ . This is shown below.

	$\Delta F^\circ_{298.2}$ (kcal.)			
$\text{H(g)} \longrightarrow \text{H}^+(\text{g}) + e$	..	..	..	$312.9$
$\text{S} \longrightarrow \text{S}^-(\text{aq.})$	..	..	..	$-320.0$
$\text{S(g)} + 2e \longrightarrow \text{S}^-(\text{g})$	..	..	..	$79.0$
$\text{H}^+(\text{g}) \longrightarrow \text{H}^+(\text{aq.})$	..	..	..	$-257.3$
$\text{SH(g)} \longrightarrow \text{S(g)} + \text{H(g)}$	..	..	..	$80.9^*$
$\text{SH}^-(\text{g}) \longrightarrow \text{SH(g)} + e$	..	..	..	$61.0^\dagger$
$\text{S}^-(\text{aq.}) + \text{H}^+(\text{aq.}) \longrightarrow \text{SH}^-(\text{aq.})$	..	..	..	$-20.5$
<hr/>				
$\text{SH}^-(\text{g}) \longrightarrow \text{SH}^-(\text{aq.})$	..	..	..	$-64$

The value of  $\Delta F_h^\circ$  for the  $\text{SH}^-$  ion is thus  $-64$  kcal./g. ion.

In this and the preceding calculation, the most inaccurate data are probably those of the electron affinities of the radicals.

## V. THE FREE ENERGY OF ION/WATER-MOLECULE INTERACTION

### *The $\chi$ -Potential of Water*

The real free energies of hydration listed in Table 1 differ from the free energies of ion/water-molecule interaction by the free energy of transport of the ion across the gas/solution interface. The difference may be calculated when

\* It is assumed that the entropy of  $\text{SH(g)}$  at  $298.2^\circ \text{K.}$  is  $45.5$  cal./deg., by analogy with the difference between the entropies of  $\text{OH(g)}$  and  $\text{H(g)}$ .

† West(39).

the  $\chi$ -potential at the free surface of the solution is known: the relationship is given by

$$\Delta F_h^\circ = \Delta F_h^{\circ'} - z\chi(\text{soln.})$$

where  $\Delta F_h^{\circ'}$  is the free energy of ion/water-molecule interaction (the "chemical" hydration energy),  $\chi(\text{soln.})$  the surface potential of the salt solution of unit activity, and  $z$  the valency of the ion.

Calculation here is facilitated by the discovery of Klei<sup>8</sup> and Lange<sup>(8)</sup> that up to relatively high concentrations the surface potential of a salt solution may be equated with reasonable accuracy to that of pure water. It may be assumed, then, that the real free energy of hydration differs from the energy of ion/water-molecule interaction by the quantity  $z\chi(\text{H}_2\text{O})$  where  $\chi(\text{H}_2\text{O})$  is the surface potential of water at the temperature of measurement. Unfortunately an accurate value of  $\chi(\text{H}_2\text{O})$  is not known independently. We may, however, calculate a value for  $\chi(\text{H}_2\text{O})$  from the difference between our values for  $\Delta F_h^\circ$  for a series of ions, and the values of Latimer, Pitzer, and Slansky<sup>(11)</sup> for  $\Delta F_h^{\circ'}$  for these ions. In Table 4 this calculation is shown for the alkali metal and halide ions.\*

TABLE 4  
THE  $\chi$ -POTENTIAL OF WATER (298.2° K.)

Ion		$-\Delta F_h^\circ$ (kcal.)	$-\Delta F_h^{\circ'}$ (kcal.)	$\chi(\text{H}_2\text{O})$ (volt)
Cs <sup>+</sup>	..	68.8	60.8	-0.35
Rb <sup>+</sup>	..	74.9	67.5	-0.32
K <sup>+</sup>	..	78.6	73.5	-0.22
Na <sup>+</sup>	..	96.1	89.7	-0.28
Li <sup>+</sup>	..	120.1	114.6	-0.24
I <sup>-</sup>	..	62.8	70.0	-0.31
Br <sup>-</sup>	..	70.8	78.0	-0.31
Cl <sup>-</sup>	..	77.2	84.2	-0.30
F <sup>-</sup>	..	106.7	113.9	-0.31

The mean value of  $\chi(\text{H}_2\text{O})$  is -0.30 v. This may be compared with the value of -0.5 v. calculated by Verwey<sup>(13)</sup> in a similar fashion.

The negative value of  $\chi(\text{H}_2\text{O})$  corresponds with an orientation of the surface dipoles such that the positive poles are turned towards the gas phase. The opposite conclusion—viz. that the oxygen poles are turned towards the gas

\* The values of Latimer, Pitzer, and Slansky<sup>(11)</sup> for the free energies of ion/water-molecule interaction of the alkali metal and halide ions are in fairly good agreement with those calculated by Verwey<sup>(13)</sup>. The figures of Latimer, Pitzer, and Slansky are employed in the calculation which follows, since the crystal energy data of Verwey and de Boer<sup>(40)</sup> which were employed by Verwey are not quite accurate (cf. Rice, 41).



phase—was reached by Frumkin(36) in comparing Latimer's values of  $\Delta F_h^\circ$  with Bernal and Fowler's(6) values for  $\Delta F_h^{\circ'}$ ; but this may be disregarded, since Verwey(13) has shown that Bernal and Fowler's calculations are greatly in error.

The question remains whether this negative value of  $\chi(\text{H}_2\text{O})$  is what might, on theoretical grounds, be expected. Little can be said about this, but the fact that in thermal agitation of the surface water molecules to produce a surface with the oxygen poles turned towards the gas phase, much more energy than for the formation of a surface with the positive poles turned towards this phase is required, is consistent with the negative value of  $\chi(\text{H}_2\text{O})$ . On the other hand, it is not in agreement with the assumption of Frumkin(37) that the more active positive poles of the water molecule are oriented towards the phase of highest dielectric constant. Provisionally, however, the mean value of  $\chi(\text{H}_2\text{O})$  may be accepted. This means that the mean difference of free energy of hydration and free energy of ion/water-molecule interaction is  $6 \cdot 9n$  kcal. for an ion of valency  $n$ .\*

## VI. APPLICABILITY OF THE MODIFIED BORN EQUATION

The applicability of the modified Born equation has been assumed in calculating the value of  $\chi(\text{H}_2\text{O})$ , since it was by this equation that Latimer's values of  $\Delta F_h^{\circ'}$  were obtained, but this may be confirmed by applying the equation to a number of other ions.

It must first be observed that the equation will only be directly applicable when the hydration energy is entirely electrostatic in character. This is not the case when an ion has a van der Waals-London potential much greater than that of the (substituted) water molecules. It is thus to be expected that with ions whose polarizability is large in comparison with the crystal radius, allowance must be made for these potentials. In order to test the Born equation directly, the hydration energies of ions with noble gas structure only will be calculated, since in these ions it can be assumed that the van der Waals contribution to the total energy is negligible.

The "chemical" hydration energies ( $\Delta F_h^{\circ'}$ ) of all ions with noble gas structure for which data are available were calculated from the real hydration energies of Table 1, with the aid of equation (15) and the value of  $\chi(\text{H}_2\text{O})$  deduced above. An examination of these figures shows that the empirical radii of Pauling(33) give much more nearly constant values of  $(r_e - r_c)$  than the Goldschmidt radii, hence these have been employed in calculating  $r_e$  for the various ions. It is found that the equations

$$-\Delta F_h^{\circ'} = \frac{Nz^2e^2}{r_c + 0 \cdot 86} (1 - 1/D) \dots\dots\dots (16a)$$

$$-\Delta F_h^{\circ'} = \frac{Nz^2e^2}{r_c + 0 \cdot 13} (1 - 1/D) \dots\dots\dots (16b)$$

yield values of  $\Delta F_h^{\circ'}$  for cations and anions respectively which are in excellent

\* The relevance of the work of Hermans(38) on the  $\zeta$ -potential at the surface of pure water, which is also negative and of the order of  $0 \cdot 1$  v., to this problem has been discussed by Verwey(13).

agreement with those calculated from the data of Table 1.\* A comparison of the two sets of data is shown in Figure 1.

By adding a term allowing for the van der Waals contribution to the total hydration energy, it should be possible to obtain a more general modification of Born's equation, applicable to ions other than noble gas structure. This problem is now being investigated.

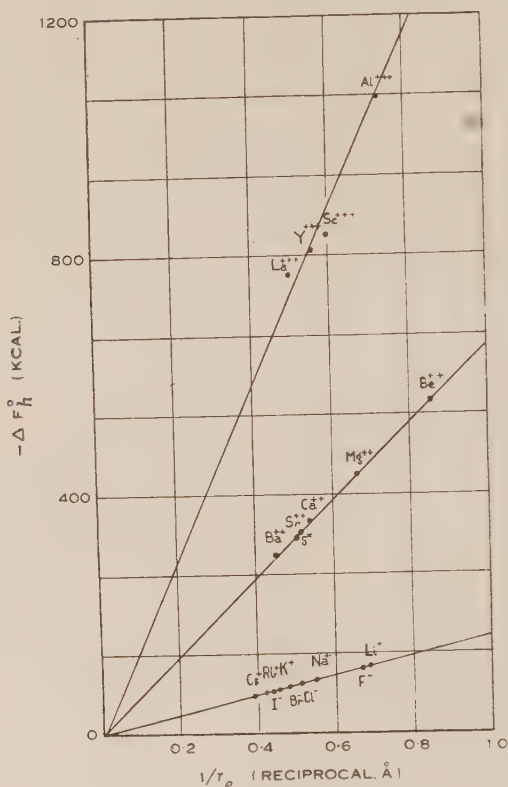


Fig. 1.—Real free energies of hydration of gaseous ions of noble gas structure as a function of effective radius  $r_e$ . Drawn-out lines are plots according to equation (16), assuming  $\chi(\text{H}_2\text{O})$  to be  $-0.30\text{v}$ .

## VII. ACKNOWLEDGMENT

The author is indebted to Dr. T. Iredale of the Department of Chemistry, University of Sydney, for critical discussion of several issues raised in this paper, which was written during tenure of the Dunlop Research Scholarship.

\* Cartledge(24) has attempted to calculate the free energy of hydration of  $\text{Hg}^+(\text{g})$  by multiplying the value for the  $\text{Ag}^+(\text{g})$  ion by the ratio of the radii of  $\text{Hg}^+$  (estimated) and  $\text{Ag}^+$ . It is highly improbable, however, that the non-electrostatic part of the total hydration energy of these ions is directly proportional to the ionic radius.

## VIII. REFERENCES

- (1) FAJANS, K.—*Ber. physik. Ges.* **21** : 549, 709 (1919).
- (2) BORN, M.—*Z. Physik.* **1** : 45 (1920).
- (3) LATIMER, W. M.—*J. Amer. Chem. Soc.* **48** : 1234 (1926).
- (4) WEBB, T. J.—*Ibid.* **48** : 2589 (1926).
- (5) ARKEL, A. E. VAN, and BOER, J. H. DE—"Chemische Bindung", p. 220 (1930).
- (6) BERNAL, J. D., and FOWLER, R. H.—*J. Chem. Phys.* **1** : 515 (1933).
- (7) VOET, A.—*Trans. Faraday Soc.* **32** : 1301 (1936).
- (8) KLEIN, O., and LANGE, E.—*Z. Elektrochem.* **43** : 570 (1937).
- (9) KLEIN, O., and LANGE, E.—*Ibid.* **44** : 562 (1938).
- (10) ELEY, D. D., and EVANS, M. G.—*Trans. Faraday Soc.* **34** : 1093 (1938).
- (11) LATIMER, W. M., PITZER, K. S., and SLANSKY, C. M.—*J. Chem. Phys.* **7** : 108 (1939).
- (12) VERWEY, E. J. W.—*Chem. Weekbl.* **37** : 530 (1940).
- (13) VERWEY, E. J. W.—*Rec. Trav. Chim. Pays-Bas* **61** : 127, 564 (1942).
- (14) LANGE, E.—*Handb. Exp. Physik.* **12** : 265 (1933).
- (15) FRUMKIN, A.—*Ergeb. Exakt Naturw.* **7** : 258 (1928).
- (16) CHALMERS, J. A.—*Phil. Mag.* **33** : 399 (1942).
- (17) BURGERS, W. G.—*Chem. Weekbl.* **39** : 198, 215 (1942).
- (18) "Internationale Tabellen zur Bestimmung von Kristallstrukturen", Vol. 2. (Berlin, 1935.)
- (19) LATIMER, W. M.—"The Oxidation States of the Elements and their Potentials in Aqueous Solution." (Prentice-Hall: New York, 1938.)
- (20) BICHOWSKY, F. R., and ROSSINI, F. D.—"The Thermochemistry of the Chemical Substances." (Reinhold: New York, 1936.)
- (21) HERZBERG, G.—"Atomic Spectra and Atomic Structure", p. 200. (Prentice-Hall: New York, 1937.)
- (22) LEWIS, G. N., and RANDALL, M.—"Thermodynamics and the Free Energy of Chemical Substances." (McGraw-Hill: New York, 1923.)
- (23) KIREEV, V.—*Acta Physiochimica U.R.S.S.* **20** : 908 (1945).
- (24) CARTLEDGE, G. H.—*J. Amer. Chem. Soc.* **63** : 906 (1941).
- (25) HUGGINS, M.—*J. Chem. Phys.* **5** : 143 (1937).
- (26) LATIMER, W. M., PITZER, K. S., and SMITH, W.—*J. Amer. Chem. Soc.* **60** : 1829 (1938).
- (27) STEPHENSON, C. C.—*Ibid.* **68** : 721 (1946).
- (28) TEMPLETON, D. H., WATT, G. W., and GARNER, C. S.—*Ibid.* **65** : 1608 (1943).
- (29) GOLDSCHMIDT, V. M.—*Nature* **157** : 192 (1946).
- (30) KAPUSTINSKII, A. F.—*J. Phys. Chem. U.S.S.R.* **5** : 73 (1939).
- (31) SHERMAN, J.—*Chem. Rev.* **11** : 146 (1932).
- (32) JONES, G., and COLVIN, J. H.—*J. Amer. Chem. Soc.* **66** : 1573 (1944).
- (33) PAULING, L.—"The Nature of the Chemical Bond." (Cornell University Press, 1939.)
- (34) GARRICK, F. J.—*Phil. Mag.* (7) **8** : 102 (1929).
- (35) BAUGHAN, E. C.—*J. Chem. Soc.* **1940** : 1403 (1940).
- (36) FRUMKIN, A.—*J. Chem. Phys.* **7** : 552 (1939).
- (37) FRUMKIN, A.—*Z. Phys. Chem.* **111** : 190 (1924).
- (38) HERMANS, J. J.—*Rec. Trav. Chim. Pays-Bas* **60** : 747 (1941).
- (39) WEST, C.—*J. Phys. Chem.* **39** : 493 (1935).
- (40) VERWEY, E. J. W., and BOER, J. H. DE—*Rec. Trav. Chim.* **55** : 451 (1936).
- (41) RICE, O. K.—"Electronic Structure and Chemical Binding." (McGraw-Hill: New York, 1940.)





# INDEX

	PAGE		PAGE
Abbey, R. L., and Barlow, G. E.		Boas, W.— <i>See</i> Honeycombe,	
The Velocity of Sound in		R. W. K.	
Gases .. .. .	175	Bolton, J. G., and Stanley, G. J.	
Absorption Spectrum (Visible)		Observations on the Variable	
of Bromine in Solution ..	472	Source of Cosmic Radio Fre-	
Alkylene Trithiocarbonates ..	236	quency Radiation in the Con-	
Alloy Containing Two Phases,		stellation of Cygnus.. ..	58
Deformation and Crystalliza-		Brass, $\beta$ , Disordering of, by	
tion of .. .. .	70	Cold Work .. .. .	190
Alunite, Synthetic and Natural,		Bromine in Solution, Visible	
Thermal Decomposition of ..	343	Absorption Spectrum of ..	472
Aromatic Ketimines, N-Acyl		Brown, G. M.— <i>See</i> Banfield,	
Derivatives of .. .. .	330	J. E., <i>et al.</i>	
Atmospheres, Hydrogen ..			
275, 289, 305, 360			
Atmospheric Pressure Changes	41		
Atomic States, Lower and			
Upper, Populations of	275, 289		
Bailey, V. A.—			
Plane Waves in an Ionized		Carbonyl Group in Thiolognin,	
Gas with Static Electric and		Presence of .. .. .	241
Magnetic Fields Present ..	351	Caro, D. E., Law, P. G., and	
B.A.L., Attempted Syntheses of	236	Rathgeber, H. D.—	
Banfield, J. E., Brown, G. M.,		Cosmic Ray Measurements	
Davey, F. H., Davies W.,		made during the Cruise of	
and Ramsay, T. H.—		H.M.A.S. <i>Wyatt Earp</i> 1947-48	261
N - Acyl Derivatives of		Champion, K. S. W.— <i>See</i>	
Aromatic Ketimines.. ..	330	Somerville, J. M., and Bigg,	
Barlow, G. E.— <i>See</i> Abbey, R. L.		E. K.	
Bayliss, N. S., Cole, A. R. H.,		“Chapman Region”, Influence	
and Green, B. G.—		of Vertical Ionic Drift on ..	423
The Visible Absorption		Chromosphere, Solar .. ..	360
Spectrum of Bromine in		Cole, A. R. H.— <i>See</i> Bayliss,	
Solution .. .. .	472	N. S., and Green, B. G.	
Bayliss, N. S., Cowley, J. M.,		Cosmic Radio Frequency Radia-	
Farrant, J. L., and Miles,		tion in Cygnus .. .. .	58
G. L.—		Cosmic Ray Measurements made	
The Thermal Decomposition		during the Cruise of H.M.A.S.	
of Synthetic and Natural		<i>Wyatt Earp</i> 1947-48.. ..	261
Alunite: An Investigation by		Cowley, J. M.— <i>See</i> Bayliss,	
X-ray Diffraction, Electron		N. S., <i>et al.</i>	
Diffraction, and Electron			
Microscope Methods ..	343	Culvenor, C. C. J., and	
Bigg, E. K.— <i>See</i> Somerville,		Davies, W.—	
J. M., and Champion, K. S. W.		Alkylene Trithiocarbonates	
		and Attempted Syntheses of	
		2, 3 - Dimercaptopropanol	
		(B.A.L.) .. .. .	236

	PAGE		PAGE
Curtis, R. G., and Hatt, H. A.— Equilibria in Furfural-Water Systems under Increased Pressure and the Influence of Added Salts upon the Mutual Solubilities of Furfural and Water .. .. .	213	Furfural Water Systems under Increased Pressure, Equilibria in .. .. .	213
Cygnus, Cosmic Radio Fre- quency Radiation in ..	58	Gases, Velocity of Sound in ..	175
Cylinder, Local Isotropy in Turbulent Wake of ..	161	Gilby, A. R., and Heymann, E. The Rate of Evaporation of Water through Duplex Films	197
Cylinders, Coaxial, Glow Dis- charge between .. ..	400	Giovanelli, R. G.— Hydrogen Atmospheres in the Absence of Thermodynamic Equilibrium	
		I. The Populations of the Upper Atomic States ..	275
Darby, J. F., and Swan, J. B.— The Angular Distribution of D-D Neutrons Scattered by Deuterons .. ..	18	II. The Populations of the Lower Atomic States ..	289
Davey, F. H.— <i>See</i> Banfield, J. E., <i>et al.</i>		III. The Emission of Radia- tion .. ..	305
Davies, W.— <i>See</i> Banfield, J. E., <i>et al.</i>		IV. The Solar Chromosphere	360
Davies, W.— <i>See</i> Culvenor, C. C. J.		Glow Discharge between Coaxial Cylinders in the Presence of a Non - Homogeneous Axial Magnetic Field .. ..	400
Derivatives of Aromatic Ketimines, N-Acyl .. ..	330	Godfrey, G. H.— Diffraction of Light from Sources of Finite Dimensions	1
Deuterons, Angular Distribution of D-D Neutrons Scattered by	18	Grant, Alison M.— <i>See</i> Hopper, V. D.	
D-D Neutrons Scattered by Deuterons .. ..	18	Green, B. G.— <i>See</i> Bayliss, N. S., and Cole, A. H. R.	
Diffraction, Electron and X-ray, in Alunite .. ..	343	Gregory, J. N., and Newing, Marjorie J.— Lubrication of Metal Surfaces by Silicone Films .. ..	85
Diffraction of Light from Sources of Finite Dimensions	1	Hatt, H. H.— <i>See</i> Curtis, R. G.	
2, 3 - Dimercaptopropanol (B.A.L.), Attempted Syn- theses of .. ..	236	Heymann, E.— <i>See</i> Gilby, A. R.	
Duplex Films, Rate of Evapora- tion of Water through ..	197	High Polymer Solutions, Statis- tical Mechanics of .. ..	319
		Honeycombe, R. W. K., and Boas, W.— Deformation and Recrystal- lization of an Alloy Contain- ing Two Phases .. ..	70
Electrons and Positrons in Gold, Scattering of .. ..	249	The Disordering of $\beta$ Brass by Cold Work .. ..	190
Electronic Charge and Oil Drop Method .. ..	369	Hopper, V. D.— The Electronic Charge and the Oil Drop Method ..	369
<i>Eucalyptus regnans</i> , Lignin of ..	112, 241	Hopper, V. D., and Grant, Alison, M.— The Influence of a Horizontal Wall on the Motion of a Falling Oil Drop .. ..	28
Farrant, J. L.— <i>See</i> Bayliss, N. S., <i>et al.</i>			
Free Energies of Hydration of Gaseous Ions .. ..	480		

	PAGE		PAGE
Howard, W.— <i>See</i> Wittrick, W. H.		Light from Sources of Finite Dimensions. Diffraction of ..	1
Hush, N. S.—		Lignin of <i>Eucalyptus regnans</i> ..	112, 241
The Free Energies of Hydration of Gaseous Ions ..	480	Lubrication of Metal Surfaces by Silicone Films .. ..	85
Hydration of Gaseous Ions, Free Energies of .. ..	480	Lugg, J. W. H., and Overell, B. T.—	
Hydrogen Atmospheres in Absence of Thermodynamic Equilibrium 275, 289, 305, 360		“ One- ” and “ Two-dimensional ” Partition Chromatographic Separations of Organic Acids on an Inert Sheet Support .. ..	98
Hydroxyl Groups in Thiolignin, Nature of .. ..	241	Merewether, J. W. T.—	
Interferometer, New .. ..	464	Studies of the Lignin of <i>Eucalyptus regnans</i> . II. The nature of the Hydroxyl Groups and the Presence of the Carbonyl Group in Thiolignin .. ..	241
Ionic Drift, Vertical, Influence of, on “ Chapman Region ” ..	423	Merewether, J. W. T.— <i>See</i> Lahey, F. N.	
Ionized Gas, Plane Waves in ..	351	Metal Surfaces, Lubrication of, by Silicone Films .. ..	85
Ions, Gaseous, Free Energies of Hydration of .. ..	480	Miles, G. L.— <i>See</i> Bayliss, N. S., <i>et al.</i>	
Isotropy in the Turbulent Wake of a Cylinder, Local .. ..	161	Miller, A. R.—	
James, R. W.—		Statistical Mechanics of High Polymer Solutions .. ..	319
Translation and Development in Two-Dimensional Fields with Special Reference to Pressure Variations .. ..	412	N - Acyl Derivatives of Aromatic Ketimines .. ..	330
Johnson, E. R., and Scholes, J. F. M.—		Neutrons, D-D, Scattered by Deuterons .. ..	18
A New Interferometer .. ..	464	Newing, Marjorie J.— <i>See</i> Gregory, J. N.	
Kerr, F. J.—		Oil Drop Method, Electronic Charge and the .. ..	369
Radio Superrefraction in the Coastal Regions of Australia ..	433	Oil Drop, Influence of Horizontal Wall on Motion of ..	28
Ketimines, N-Acyl Derivatives of Aromatic .. ..	330	Overell, B. T. — <i>See</i> Lugg, J. W. H.	
Kirkpatrick, C. B.—		Organic Acids, Partition Chromatographic Separations of .. ..	98
The Influence of Vertical Ionic Drift on a “ Chapman Region ” .. ..	423	Partition Chromatographic Separations of Organic Acids .. ..	98
Lahey, F. N., and Merewether, J. W. T.—		Pivots, Symmetrical Crossed Flexure, Theory of .. ..	121
Studies on the Lignin of <i>Eucalyptus regnans</i> . I. Thiolignin .. ..	112		
Lasich, W. B.—			
The Scattering of Fast Positrons and Electrons in Gold (for the Range 0.39–1.1 Mev.) .. ..	249		
Law, P. G.— <i>See</i> Caro, D. E., and Rathgeber, H. D.			

	PAGE		PAGE
Plane Waves in an Ionized Gas with Static Electric and Mag- netic Fields Present.. ..	351	Somerville, J. M., Champion, K. S. W., and Bigg, E. K.— Some Aspects of the Glow Discharge Between Coaxial Cylinders in the Presence of a Non-Homogeneous Axial Magnetic Field .. ..	400
Positrons in Gold, Scattering of Electrons and .. ..	249	Sound in Gases, Velocity of ..	175
Pressure Changes, Atmospheric	41	Spectrum of Bromine in Solu- tion, Visible Absorption ..	472
Pressure Increased, Equilibria in Furfural-Water Systems under .. ..	213	Stanley, G. J.— <i>See</i> Bolton, J. G.	
Pressure Variations in Two- Dimensional Fields .. ..	412	Stress Distribution, Relaxation Methods Applied to ..	135
Priestley, C. H. B.— Atmospheric Pressure Changes: The Importance of Deviations from the Balanced (Gradient) Wind .. ..	41	Swan, J. B.— <i>See</i> Darby, J. F.	
Radiation, Cosmic Frequency, Variable Source of, in Cygnus	58	Thermal Decomposition of Syn- thetic and Natural Alunite..	343
Radio Superrefraction in the Coastal Regions of Australia	443	Thermodynamic Equilibrium, Hydrogen Atmospheres in the Absence of .. 275, 289, 305, 360	
Ramsay, T. H.— <i>See</i> Banfield, J. E., <i>et al.</i>		Thiolignin .. ..	112, 241
Rathgeber, H. D.— <i>See</i> Caro, D. E., and Law, P. G.		Townsend, A. A.— Local Isotropy in the Tur- bulent Wake of a Cylinder	161
Relaxation Methods Applied to Two-Dimensional Stress Dis- tribution .. ..	135	Trithiocarbonates, Alkylene ..	236
Salts, Influence of, on Mutual Solubilities of Furfural and Water .. ..	213	Wake of Cylinder, Turbulent, Local Isotropy in .. ..	161
Scholes, J. F. M.— <i>See</i> Johnson, E. R.		Waves, Plane, in an Ionized Gas	351
Silicone Films, Lubrication of Metal Surfaces by .. ..	85	Wittrick, W. H.— The Theory of Symmetrical Crossed Flexure Pivots ..	121
Solar Chromosphere .. ..	360	Wittrick, W. H., and Howard, W.— Relaxation Methods Applied to Two Problems of Two- Dimensional Stress Distribu- tion Involving Mixed Boundary Conditions ..	135
Solubilities, Mutual, of Furfural and Water, Influence of Added Salts on .. ..	213	Wyatt Earp, H.M.A.S., Cosmic Ray Measurements made during the Cruise of ..	261





# DATE DUE

DEC SEP 2 1976







3 8198 304 260 605

UNIVERSITY OF ILLINOIS AT CHICAGO

4





